

ЯКОВИН СЕРГІЙ

Івано-Франківський національний технічний університет нафти і газу

<https://orcid.org/0000-0002-3335-2892>e-mail: syakovyn@gmail.com**МЕЛЬНИЧУК СТЕПАН**

Івано-Франківський національний технічний університет нафти і газу

<https://orcid.org/0000-0002-6973-4235>e-mail: stenni@ukr.net

ОБМЕЖЕННЯ МНОЖИНИ МОЖЛИВИХ РІШЕНЬ ДЛЯ ДИСКРЕТНОГО ПЕРЦЕПТРОНА ЗІ ЗМІЩЕНИМИ СИНАПТИЧНИМИ СИГНАЛАМИ ПРИ ЕМУЛЯЦІЇ БІНАРНИХ ЛОГІЧНИХ ФУНКЦІЙ

Традиційно перцептрони є основним компонентом для багатьох сучасних моделей нейронних мереж. Відомо, що базовий дискретний перцептрон придатний лише для лінійно роздільних задач, фактично теоретичні дослідження, у своїй більшості, спрямовані на дослідження багатошарових перцептронів (MLP) а також глибоких нейронних мереж, які здатні працювати з нелінійними задачами. Тут основним напрямком є вдосконалення процесу машинного навчання, зокрема в частині пошуку прийнятних рішень на обмежених даних. Теоретичний аналіз дозволяє зрозуміти, скільки навчальних зразків потрібно для ефективного навчання і наскільки добре модель буде працювати на нових наборах даних. Перцептрон зазвичай розглядається як порівняно проста модель, що є наближеною інформаційною аналогією біологічного нейрону. Фактично цифрова модель перцептрону базується на основі кореляційного підходу щодо опрацювання вхідних сигналів з подальшим використанням функції активації. Іншими словами агрегація вхідних сигналів перцептрона реалізується як сума добутоків цих сигналів з відповідними ваговими коефіцієнтами. Функції автокореляції в такому випадку відіграють ключову роль, дозволяючи встановити залежність між вхідними сигналами синапсів та цільовими значеннями, які система намагається ідентифікувати та класифікувати. Зокрема, кореляційний підхід дає змогу визначити, які вхідні сигнали синапсів формують найбільший вплив на результат опрацювання, що покращує навчання моделі. Однак, кореляція ґрунтується на використанні операції множення, що суттєво ускладнює схемні рішення перцептронних структур.

Доцільно зазначити, що часто поза увагою залишається такий напрямок розвитку як вдосконалення та розширення функціоналу структури перцептрону, зокрема в частині переходу до дискретного базису. Вивчення обмежень перцептронних структур та можливих способів їх подолання, зокрема застосування нових підходів до структурної організації, залучення різних функцій активації а також через використання багатошарових структур, може дати можливість створювати більш ефективні інструменти для вирішення задач класифікації та прогнозування.

Ключові слова: перцептрон, бінарні сигнали, логічні функції, імовірнісні оцінки, комп'ютерні компоненти

YAKOVYN SERHIY**MELNYCHUK STEPAN**

Ivano-Frankivsk National Technical University of Oil and Gas

LIMITATIONS ON THE SET OF POSSIBLE SOLUTIONS FOR A DISCRETE PERCEPTRON WITH SHIFTED SYNAPTIC SIGNALS IN EMULATION OF BINARY LOGIC FUNCTIONS

Traditionally, perceptrons have been a core component of many modern neural network models. It is well known that the basic discrete perceptron is only suitable for linearly separable tasks; in fact, most theoretical research focuses on studying multilayer perceptrons (MLPs) as well as deep neural networks capable of solving nonlinear problems. The primary direction here is the improvement of the machine learning process, particularly in terms of finding acceptable solutions with limited data. Theoretical analysis makes it possible to understand how many training samples are required for effective learning and how well the model will perform on new datasets. The perceptron is generally regarded as a relatively simple model, which serves as an approximate informational analogy to a biological neuron. In essence, the digital model of the perceptron is based on a correlation approach to processing input signals, followed by the use of an activation function. In other words, the aggregation of input signals in a perceptron is implemented as the sum of the products of these signals and their respective weight coefficients.

In this context, autocorrelation functions play a key role, enabling the establishment of dependencies between the input signals of synapses and the target values that the system seeks to identify and classify. Specifically, the correlation approach allows for identifying which synapse input signals exert the greatest influence on the processing outcome, thereby enhancing the model's training process. However, correlation relies on the use of multiplication operations, which significantly complicates the hardware implementation of perceptron structures.

It is worth noting that the development aimed at enhancing and expanding the functionality of the perceptron structure, especially with regard to transitioning to a discrete basis, is often overlooked. Studying the limitations of perceptron structures and possible ways to overcome them, including the application of new approaches to structural organization, the adoption of various activation functions as well as the use of multilayer structures, may provide the opportunity to create more effective tools for solving classification and forecasting tasks.

Keywords: perceptron, binary signals, logical functions, probabilistic estimates, computer components.

Стаття надійшла до редакції / Received 11.04.2025

Прийнята до друку / Accepted 26.04.2025

Вступ

Сучасні реалізації нейронних мереж, включаючи багатошарові структури, значною мірою ґрунтуються на фундаментальних підходах, розроблених у межах досліджень перцептрону. Класичний перцептрон став відправною точкою, з якої сформувалися сучасні нейронні мережі. Тому кожен підхід, що покращує розуміння фундаментальних механізмів навчання залишається актуальним. Перцептрон традиційно вважають відносно простою моделлю, яка ґрунтується на кореляційному опрацюванні, застосування нових ідей, зокрема фрактальної структури даних, використання ймовірнісних оцінок чи переходу до дискретного базису, дає можливість пояснити, як топологія зв'язків або структура вхідних даних впливає на швидкість ефективність і стійкість навчання. До ключових проблеми, що виникають при навчанні дискретних перцептронів з типовою структурою (див.рис.1а) мається на увазі перцептронів із пороговою чи східчастою функцією активації, що формує "дискретний" вихід, слід віднести [1-3]:

- непридатність до вирішення нелінійно роздільних задач, тобто перцептрон здатний коректно навчитися лише у випадку, коли вибірка навчальних даних є лінійно віддільною, зокрема такі логічні функції як AND, OR, NAND, NOR та інші, що є лінійно роздільними, нескладно відтворюються за допомогою простих перцептронів, однак, якщо дана задача не є лінійно роздільною, наприклад логічна функція XOR, то простий дискретний перцептрон не може знайти розв'язок у принципі;

- відсутність похідної та "сигналу" для градієнтних методів, тобто для порогової активації похідна не визначена в точці порогу, отже звичні методи (зокрема градієнтного спуску) не застосовуються безпосередньо оскільки неможливо обчислити градієнтну похибку в традиційному вигляді. В такій ситуації, дискретний перцептрон можна навчати за допомогою правил перцептронної (Hebbian-like) корекції - коригування ваг лише за наявності помилки. Недоліком такого підходу є те, що процес навчання може довго коливатися навколо коректного рішення, зокрема для випадків коли вибірка майже лінійно роздільна або містить шум;

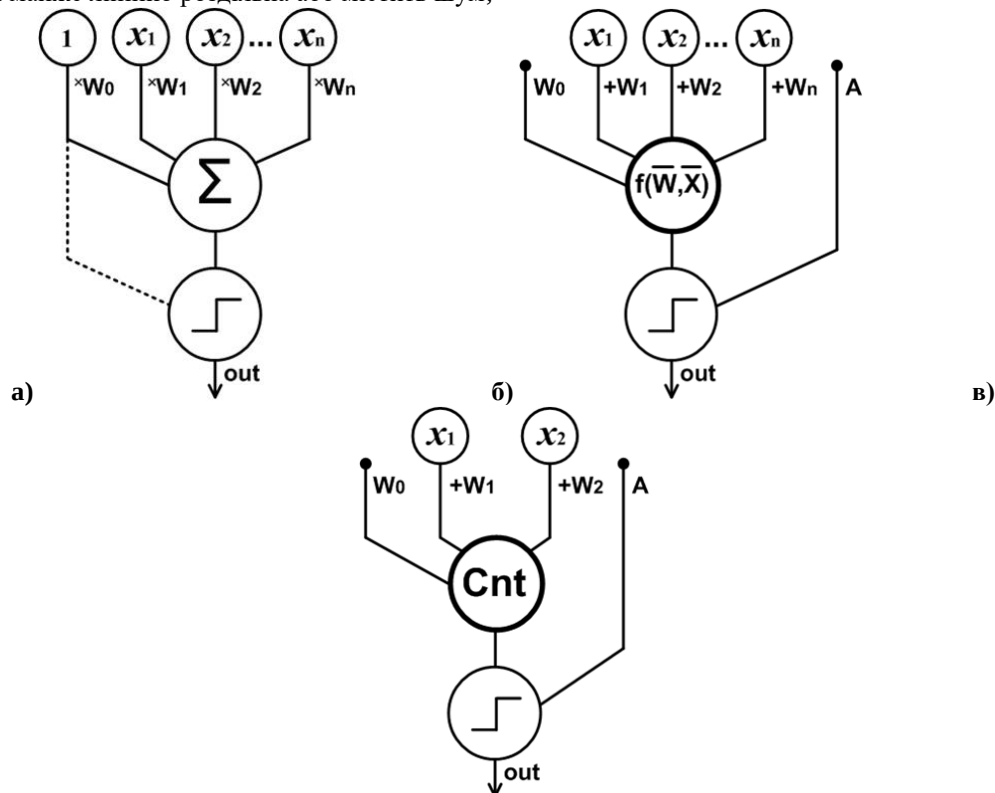


Рис. 1. Структура перцептрону на основі: а) суматора (типова) на n входів; б) ймовірнісних оцінок зміщених сигналів синапсів на n входів; в) функції Cnt - підрахунку кількості різних станів зміщених синаптичних сигналів на 2 бінарні входи

- надмірна чутливість до зашумленості вхідних даних, оскільки вихід перцептрону дискретний і змінюється стрибкоподібно то навіть незначні коливання на вході в околі порогу можуть змінити результат на протилежний. В такій ситуації використовують методи згладжування вхідних сигналів чи використовують інші функції активації;

- низька роздільна/виразна здатність, оскільки дискретний перцептрон є лінійним класифікатором то він не може описати нелінійні розділові поверхні без доповнення його додатковими шарами чи особливими методами трансформації ознак;

- проблема зупинки навчання, якщо набори вхідних даних не є ідеально лінійно роздільними чи містять протиріччя, алгоритм навчання може бути не збіжним. Це може проявлятися в ситуації безкінечного навчання (постійне коригування ваг у різні боки після кожної помилки) без гарантії зменшення помилок до нуля, або в осциляціях навколо часткового рішення;

- потреба в обережному виборі гіперпараметрів, оскільки швидкість навчання (learningrate) та

ініціалізація вагових коефіцієнтів суттєво впливають на часову складність навчання то некоректно обраний крок оновлення може спричиняти зайві стрибки ваг і тривале чи навіть нескінченне навчання.

Таким чином, основні проблеми можна звести до того, що: дискретна активація ускладнює використання градієнтних методів, модель від природи має низьку роздільну/виразну здатність, а зашумлені або нелінійно роздільні дані роблять навчання проблематичним або неможливим.

В такій ситуації доцільно згадати реалізацію перцептронів (див.рис.1б), що ґрунтується на використанні імовірнісних характеристик, зокрема оцінюванні міри хаотичності а саме інформаційної ентропії вхідних дискретних сигналів синапсів [4]. Фактично тут запропоновано перехід до дискретного базису, хоча структура перцептронів залишається незмінною, однак опрацювання зважених (перемножених з відповідними ваговими коефіцієнтами) синаптичних сигналів, які формують множину з обмеженою кількістю дискретних станів, передбачає оцінювання інформаційної ентропії такої множини з подальшим залученням порогової функції активації. Такий підхід є нечутливим до абсолютних значень сигналів синапсів а враховує тільки імовірності їх появи. До проблемних моментів згаданої структури слід віднести те, що формула ентропії містить функцію розрахунку логарифму і це призводить до додаткових обчислювальних затрат а також відсутність досліджень щодо особливостей навчання таких перцептронних структур.

Таким чином, перцептрон, як базова ланка класичних нейронних мереж, можна розглядати у "статистико-механічних" категоріях, зокрема: розподіл вагових коефіцієнтів, динаміку навчання та вплив випадковості в процесі тренування. В статті [5] розглянуто методи, які дозволяють оцінювати зміни в параметрах та структуру просторів станів, що можуть бути адаптовані для аналізу перцептронів: як складність чи хаотичність даних відображається на здатності нейронної мережі навчатися. Знання щодо критичних режимів або фазових переходів (keytransitions) у складних системах дають можливість, коли алгоритм навчання перцептронів може "застрягати" а коли він здатний до ефективної узагальнюючої здатності. Тобто, якщо можна виявити певні "точки фазового переходу", зокрема, межу між простими та складними для навчання наборами даних, це може розширити можливості для покращення алгоритмів тренування.

Загалом, актуальність досліджень у цій сфері полягає у розширенні інструментарію для аналізу та вдосконалення процедури навчання перцептронів, як компоненту нейромережевої структури, через вдосконалення наявних і залучення нових, зокрема структурних, рішень. Крім того, це розширює можливості для створення більш ефективних алгоритмів, оптимальних стратегій навчання а також розуміння механізмів узагальнення і, як наслідок, сприяє розвитку інструментальних засобів, що забезпечують ефективну платформу для розвитку нейронних мереж. Таким чином, навіть такі базові компоненти як перцептрон залишаються фундаментально важливими для наукових досліджень і практичних реалізацій, особливо на основі сучасних апаратних платформ, зокрема програмованих логічних інтегральних схем.

Метою роботи є: дослідження, з метою зменшення, множини можливих рішень для дискретного перцептрона із сигналами синапсів, що зазнали зміщення, для емуляції бінарних логічних функцій на два інформаційні входи.

Результат дослідження

Доцільно зазначити, що аналіз дискретних перцептронів і досі лежить в основі розуміння лінійної роздільності та обчислювальної спроможності нейронних мереж а класичне правило навчання перцептронів має сформульовані докази збіжності лише для лінійно віддільних випадків.

Реалізація перцептронів, що використовує імовірнісні оцінки (див.рис.1б), зокрема інформаційну ентропію вхідних дискретних сигналів синапсів [4], нескладно спростити замінивши формулу розрахунку ентропії К.Шеннона: $H = -k \sum_{i=1}^n p_i \cdot \log p_i$ яка для випадку рівномірного розподілу вхідних сигналів (даних) може бути зведена до формули С.Хартлі: $H = \log_2 S^n = n \cdot \log_2 S$. Тут доцільно зазначити, що кількість входів незмінна для всіх випадків, тому коефіцієнтом n можна знехтувати. Крім того, відмова від використання функції логарифму дозволяє суттєво зменшити обчислювальні затрати без порушення адекватності результату. В результаті запропоновано структуру перцептронів, у якій зміна стану активації (збудження) ґрунтується на оцінюванні кількості унікальних станів (значень, що не повторюються) шляхом додавання відповідних дискретних коефіцієнтів зміщення до сигналів синапсів [6].

В такій ситуації доцільно провести дослідження множини можливих рішень, зокрема на прикладі перцептронної структури (див.рис.1в), що здатна емулявати вичерпний перелік логічних функцій для 2-ох бінарних вхідних сигналів. При виборі коефіцієнтів зміщення (що можна вважати деяким аналогом ваг) перцептронів одне значення, для наочності, необхідно зафіксувати як константу, зокрема додатковий вхід $w_0=0$. Такий вибір пояснюється тим, що зміщення всіх коефіцієнтів на однакову величину не змінює кількості унікальних станів. Так для логічних функцій OR значення коефіцієнтів зміщення (w_0, w_1, w_2) для 2-ох входів перцептронів можуть бути $[0,0,0]$ або $[1,1,1]$ або $[2,2,2]$. Одним із ефективних прикладів знаходження допустимих значень коефіцієнтів зміщення, в цьому випадку, може бути використання матриці кількості унікальних станів при фіксованому значенні додаткового входу w_0 , яка формується таким чином: для кожної комбінації $(0, w_1, w_2)$ знаходиться кількість унікальних станів. Наприклад: $(0,0,0)$ дає 1 стан, $(0,1,1)$ - 2 стани.

Приклад реалізації такої матриці для 2-ох входового дискретного перцептрон, при зафіксованому значенні додаткового входу $w_0=0$ подано в табл.1. Як можна побачити отримана матриця є симетричною відносно головної діагоналі, що дозволяє скоротити, я мінімум наполовину, множину пошуку можливих коефіцієнтів зміщення, які необхідно опрацювати, при реалізації перцептроном відповідної логічної функції.

Знаходження прийнятних значень коефіцієнтів зміщення w_1 та w_2 здійснюється шляхом порівняння можливих значень з таблиці істинності обраної логічної функції, для всіх комбінацій входів x_1 та x_2 , із відповідно розташованими значеннями фрагменту ідентичної розмірності (виділено пунктиром) у таблиці кількості унікальних станів синаптичних сигналів 2-ох входового дискретного перцептрон.

Таблиця 1

Кількість унікальних станів зміщених синаптичних сигналів 2-ох входового дискретного перцептрон для коефіцієнтів w_1 та w_2 зміщення [-4,4] при зафіксованому значенні $w_0=0$

$w_1 \backslash w_2$	-4	-3	-2	-1	0	1	2	3	4
-4	2	3	3	3	2	3	3	3	3
-3	3	2	3	3	2	3	3	3	3
-2	3	3	2	3	2	3	3	3	3
-1	3	3	3	2	2	3	3	3	3
0	2	2	2	2	1	2	2	2	2
1	3	3	3	3	2	2	3	3	3
2	3	3	3	3	2	3	2	3	3
3	3	3	3	3	2	3	3	2	3
4	3	3	3	3	2	3	3	3	2

Тобто, в ході пошуку значень коефіцієнтів зміщення w_1 та w_2 , що задовольняють умовам емуляції відповідної логічної функції відповідний фрагмент таблиці істинності почергово накладається на відповідні фрагменти матриці кількості унікальних станів з метою виявлення характерних відношень між відповідними значеннями таблиці істинності та аналогічно розташованими значеннями матриці.

Іншими словами необхідно виявити функціональну відповідність між значеннями таблиці істинності логічної функції та значеннями матриці кількості унікальних станів. Результати знаходження коефіцієнтів зміщення для емуляції логічних функцій OR та NOR подано в таблиці 2.

Таблиця 2

Доступні рішення емуляції логічного компоненту OR та NOR на основі 2-ох входового дискретного перцептрон для коефіцієнтів w_1 та w_2 зміщення [-4,4] при зафіксованому значенні $w_0=0$

$w_1 \backslash w_2$	-4	-3	-2	-1	0	1	2	3	4
-4	2	3	3	3	2	3	3	3	3
-3	3	2	3	3	2	3	3	3	3
-2	3	3	2	3	2	3	3	3	3
-1	3	3	3	2	2	3	3	3	3
0	2	2	2	2	1	2	2	2	2
1	3	3	3	3	2	2	3	3	3
2	3	3	3	3	2	3	2	3	3
3	3	3	3	3	2	3	3	2	3
4	3	3	3	3	2	3	3	3	2

OR	0	1
0	0	1
1	1	1

NOR	0	1
0	1	0
1	0	0

З аналогічним підходом не складно знайти значення коефіцієнтів зміщення для логічних функцій AND, NAND, XOR та XNOR. Результати знаходження коефіцієнтів зміщення для емуляції згаданих логічних функцій подано в таблицях 3 та 4.

Таблиця 3

Доступні рішення емуляції логічного компоненту AND та NAND на основі 2-ох входового дискретного перцептрону для коефіцієнтів w_1 та w_2 зміщення $[-4,4]$ при зафіксованому значенні $w_0=0$

			w1 \ w2									
			-4	-3	-2	-1	0	1	2	3	4	
-4			2	3	3	3	2	3	3	3	3	
-3			3	2	3	3	2	3	3	3	3	
-2			3	3	2	3	2	3	3	3	3	
-1			3	3	3	2	2	3	3	3	3	
0			2	2	2	2	1	2	2	2	2	
1			3	3	3	3	2	2	3	3	3	
2			3	3	3	3	2	3	2	3	3	
3			3	3	3	3	2	3	3	2	3	
4			3	3	3	3	2	3	3	3	2	

AND	0	1
0	0	0
1	0	1

NAND	0	1
0	1	1
1	1	0

Як можна побачити, реалізація логічних функцій XOR та XNOR може бути представлена тим ж коефіцієнтами зміщень але використано різні функції активації $=$ та $\neq$ відповідно. Така ситуація пояснюється тим, що значення функцій XOR та XNOR симетричні відносно головної та побічної діагоналей (див. таблиці істинності цих функцій).

Таблиця 4

Доступні рішення емуляції логічного компоненту XOR та XNOR на основі 2-ох входового дискретного перцептрону для коефіцієнтів w_1 та w_2 зміщення $[-4,4]$ при зафіксованому значенні $w_0=0$

			00									
--	--	--	--	--	--	--	--	--	--	--	--	--

Доцільно зазначити, що при реалізації логічних функцій OR, AND, NOR та NAND є можливість маніпуляції як коефіцієнтами зміщення w_1 та w_2 так і функціями активації оскільки вони симетричні тільки відносно головної діагоналі.

Фактично, вибір функції активації, зокрема для розглянутих випадків це: $<$, $>$, $\leq$, $\geq$, $=$, $\neq$, безпосередньо впливає на оцінювання відношень між елементами матриці кількості унікальних станів, що необхідно враховувати при виборі необхідних значень коефіцієнтів зміщень.

Оскільки функції XOR та XNOR мають ідентичні рішення в матриці кількості унікальних станів (див. табл.4), то в такому випадку доцільно використовувати від'ємні значення порогу функції

активації як ознаку інверсії результату виходу перцептрону.

Якщо узагальнити прийнятні варіанти коефіцієнтів зміщення перцептрону для усіх вищерозглянутих логічних функцій OR, AND, XOR, NOR, NAND, XNOR, результат подано в табл.5, то можна побачити, що множина розв'язків, значень w_1 та w_2 при зафіксованому $w_0=0$, групується в околі головної діагоналі.

Таблиця 5

Узагальнена множина доступних рішень емуляції логічних компонентів OR, AND, XOR, NOR, NAND, XNOR на основі 2-ох входового дискретного перцептрону для коефіцієнтів w_1 та w_2 зміщення [-4,4] при зафіксованому значенні $w_0=0$

$w_1 \backslash w_2$	-4	-3	-2	-1	0	1	2	3	4
-4	2	3	3	3	2	3	3	3	3
-3	3	2	3	3	2	3	3	3	3
-2	3	3	2	3	2	3	3	3	3
-1	3	3	3	2	2	3	3	3	3
0	2	2	2	2	1	2	2	2	2
1	3	3	3	3	2	2	3	3	3
2	3	3	3	3	2	3	2	3	3
3	3	3	3	3	2	3	3	2	3
4	3	3	3	3	2	3	3	3	2

Таким чином встановлено, що для запропонованої структури 2-ох входового дискретного перцептрону прийнятні розв'язки слід шукати вздовж головної діагоналі матриці кількості унікальних станів околі, який не перевищує 35% вибірки для розглянутих значень коефіцієнтів w_1 та w_2 . З огляду на такий отримані результати можна припустити, що тренд розташування можливих розв'язків буде зберігатися і для інших логічних функцій, зокрема на більшу кількість входів.

Висновок

За результатами проведених досліджень показано можливість реалізації перцептронних структур, зокрема для дискретного базису, що здатні вирішувати задачі поза традиційно прийнятими межами лінійної роздільності.

Крім того, встановлено можливість суттєвого, до 35% від загального обсягу, обмеження простру можливих рішень, що дозволяє прискорити процес знаходження (навчання) коефіцієнтів зміщення при опрацюванні сигналів синапсів на основі оцінювання кількості їх унікальних дискретних станів. Доцільно зазначити, що розглянута структура дискретного перцептрону не потребує використання операції множення, яка замінена на операцію зміщення (реалізується шляхом додавання відповідних коефіцієнтів), що дозволяє маніпулювати імовірнісними оцінками сигналів синапсів, а також зменшує апаратних затрат на реалізацію.

Хоча багатопарові мережі з неперервними функціями активації (sigmoid, ReLU) на сучасному етапі розвитку комп'ютерних систем домінують, дискретні перцептрони є актуальними і використовуються для спеціалізованих логічних схем, апаратних реалізацій, зокрема на платформах програмованих логічних інтегральних схем (ПЛІС) FPGA, CPLD, ASIC тощо.

Література

1. Haykin, S. (2009). *Neural networks and learning machines*. Pearson Education, Limited.
2. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer.
3. Bengio, Y., Courville, A., & Goodfellow, I. (2016). *Deep learning*. MIT Press.
4. Melnychuk, S., Kuz, M., & Yakovyn, S. (2018). Emulation of logical functions NOT, AND, OR, and XOR with a perceptron implemented using an information entropy function. In *2018 14th international conference on advanced trends in radioelectronics, telecommunications and computer engineering (TCSET)*. IEEE. <https://doi.org/10.1109/tcset.2018.8336337>
5. Rosen-Zvi, M., & Kanter, I. (2001). Training a perceptron in a discrete weight space. *Physical Review E*, 64(4). <https://doi.org/10.1103/physreve.64.046109>

6. Мельничук, С. І., & Яковин, С. В. (2023). *Спосіб реалізації перцептрона на основі імовірнісних характеристик зміщених синаптичних сигналів* (Ukrainian Patent No. 126753).

References

1. Haykin, S. (2009). *Neural networks and learning machines*. Pearson Education, Limited.
2. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer.
3. Bengio, Y., Courville, A., & Goodfellow, I. (2016). *Deep learning*. MIT Press.
4. Melnychuk, S., Kuz, M., & Yakovyn, S. (2018). Emulation of logical functions NOT, AND, OR, and XOR with a perceptron implemented using an information entropy function. In *2018 14th international conference on advanced trends in radioelectronics, telecommunications and computer engineering (TCSET)*. IEEE. <https://doi.org/10.1109/tcset.2018.8336337>
5. Rosen-Zvi, M., & Kanter, I. (2001). Training a perceptron in a discrete weight space. *Physical Review E*, 64(4). <https://doi.org/10.1103/physreve.64.046109>
6. Melnychuk, S.I., & Yakovyn, S.V. (2023). *Sposib realizatsii pertseptrona na osnovi imovirnisnykh kharakterystyk zmishchenykh synaptychnykh syhnaliv* (Ukrainian Patent No. 126753).