

УДК 004.8-044.3

<https://doi.org/10.31713/vt2202516>

Турбал Ю. В. [1; ORCID ID: 0000-0002-5727-5334],

д.т.н., професор,

Кубай О. В. [1; ORCID ID: 0000-0002-2005-487X],

аспірант

¹Національний університет водного господарства та природокористування, м. Рівне

ТЕОРЕМА КОВЕРА ТА ЗАДАЧА ВІДОКРЕМЛЕННЯ МНОЖИН ЗА ДОПОМОГОЮ ПОЛІНОМІАЛЬНИХ ПОРОГОВИХ ФУНКЦІЙ (PTF – POLYNOMIAL THRESHOLD FUNCTIONS)

Відокремлення підмножин точок у евклідовому просторі – базова задача аналізу даних і нейрообчислень; її розв'язність істотно залежить від класу допустимих розділювальних поверхонь. Лінійні моделі мають обмеження (для 3 змінних лінійно роздільними є лише 104 із 256 дихотомій), що спонукає перейти до нелінійних підходів.

У статті представлений повний комбінаторний перелік усіх 256 можливих розділень вершин, точно кількісно визначено межі роздільності поліномами другого ступеня. Автори показують, що з 22 унікальних класів еквівалентності (орбіт) дихотомій, які існують відносно групи симетрії куба, лише один клас є квадратично нероздільним. Цей аналіз не лише дає вичерпну відповідь на поставлене питання, але й слугує ілюстрацією фундаментальних обмежень, властивих навіть потужним нелінійним моделям при роботі з високоструктурованими даними.

Для досягнення цієї мети автори використовують ізоморфізм між геометричною задачею розділення та алгебраїчною задачею реалізації булевих функцій. Кожна дихотомія вершин куба однозначно відповідає булевій функції трьох змінних, а питання квадратичної роздільності стає еквівалентним питанню про те, чи є така функція поліноміальною пороговою функцією другого ступеня.

Результат дає точну межу виражальної здатності квадратичних порогових елементів для задач бінарної класифікації у низьких вимірах і пояснює, чому вже перехід від лінійних до квадратичних моделей кардинально підсилює класифікаційну потужність.

Ключові слова: поліноміальна порогова функція; дихотомія куба; теорема Ковера; теорема Пойя; трюк ядра (kernel trick); квадратичні розділювачі; нейронна мережа; булева функція.

Вступ. Задача розділення певною поверхнею заданої сукупності точок у деякому просторі є фундаментальною проблемою



в таких галузях, як аналіз даних, машинне навчання та нейронні мережі [3]. Виразальна здатність розділювальної поверхні – чи то гіперплощини, чи то алгебраїчної поверхні вищого порядку – безпосередньо визначає клас задач класифікації, які можна розв'язати.

Перехід від лінійних до квадратичних розділювачів є класичним кроком у подоланні обмежень простих моделей. Добре відомо, що не всі дихотомії вершин куба є лінійно роздільними; канонічним контрприкладом є функція парності (XOR), геометрична нероздільність якої зумовлена перетином опуклих оболонок відповідних підмножин вершин. Цей факт, що історично продемонстрував межі одношарових перцептронів, потребує дослідження більш потужних, нелінійних класифікаторів. Квадратичні поверхні є природним наступним кроком у цій ієрархії складності. Кожна дихотомія вершин куба однозначно відповідає булевій функції трьох змінних, а питання квадратичної роздільності стає еквівалентним питанню про те, чи є така функція поліноміальною пороговою функцією другого ступеня.

Метою роботи є визначення точної кількості дихотомій куба, які можна реалізувати за допомогою поліноміальної порогової функції (PTF) 2 ступеня.

1. Проблема відокремлення множин і перцептронне зображення логічних функцій

Задача відокремлення множин і нейронне зображення логічних функцій є відомою проблемою в процесі конструювання нейронних мереж. Кожен нейрон є формальним пороговим елементом, що набуває одиничного значення, якщо сумарний зважений вхід більше деякого граничного значення, а в протилежному випадку – нуль:

$$y_j = \begin{cases} 1, & \sum_{j=1}^n w_{ij}x_j > Q_j, \\ 0, & \sum_{j=1}^n w_{ij}x_j \leq Q_j, \end{cases}$$

де n – кількість нейронів у сітці.

Таким чином, за заданих значень ваг і порогів, нейрон має певну числову характеристику вихідної активності для кожного можливого вектора входів. Множина вхідних векторів, при яких нейрон активний ($y_j = 1$), відділена від множини векторів, на яких нейрон пасивний ($y_j = 0$), гіперплощиною, рівняння якої таке:

$$\sum_{i=1}^n w_{ij}x_j - Q_j = 0.$$

Отже, нейрон здатний відокремити (мати різний вихід) тільки такі дві множини векторів входів, для яких є гіперплощина, що відтинає одну множину від іншої. Такі множини називають лінійно-роздільними [1]. Однак, не всі множини можна лінійно відокремити одна від одної. На рис. 1 наведено приклад, коли цього зробити не можна.

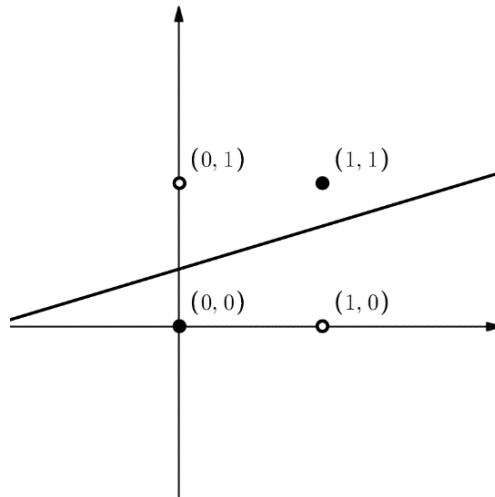


Рис. 1. Неможливість лінійного відокремлення чорних і білих точок

При зростанні кількості аргументів ситуація стає ще гіршою: кількість функцій, які можна лінійно відокремити, різко зменшується. А значить, і звужується клас функцій, що може бути реалізований нейроном (див. табл. 1) [1].

Таблиця 1

Кількість змінних N	Повна кількість можливих логічних функцій 2^{2^N}	Кількість функцій, що лінійно відокремлені	Відсоток до загальної кількості
1	4	4	100
2	16	14	87,5
3	256	104	40,6
4	65536	1882	2,9
5	4294967296	94572	0,002

Із табл. 1 бачимо, що можливість одношарового нейрона точно зображувати наперед задану логічну функцію, має значні обмеження. Для усунення таких обмежень була запропонована концепція використання кількох шарів нейронів [2]. Із збільшенням



кількості шарів зростає й складність архітектури нейронної мережі.

2. Застосування поліноміальної функції другого степеня для розділення вершин куба

Розглянемо ідею використання нелінійної функції, а саме полінома другого степеня, для розділення множини вершин одиничного куба Q_3 в тривимірному евклідовому просторі. Такий куб є опуклою оболонкою його восьми вершин. Їх можна представити набором бінарних векторів $V = \{0,1\}^3$. Опишемо деяку комбінацію вершин цього куба як дихотомію множини V , тобто, розбиття V на дві підмножини A і B такі, що $A \cup B = V$, а $A \cap B = \emptyset$. Кожен унікальний вибір підмножини вершин A визначає унікальну дихотомію, де B є її доповненням. Оскільки існує $2^{|V|} = 2^8 = 256$ можливих підмножин, потрібно розглянути 256 різних дихотомій (включаючи тривіальні випадки, коли $A = \emptyset$ або $A = V$).

В процесі розв'язування поставленої задачі виникло питання – яку ж саме квадратичну поверхню застосувати для дихотомії вершин куба? Першочергово ми розглянули еліпсоїд, який можна описати за допомогою канонічного рівняння (1):

$$\frac{(x - x_1)^2}{a} + \frac{(y - y_1)^2}{b} + \frac{(z - z_1)^2}{c} = 1. \quad (1)$$

На основі рівняння (1) ми побудували систему нерівностей (2) для вершин одиничного куба:

$$\left\{ \begin{array}{l} \frac{(1 - x_1)^2}{a} + \frac{y_1^2}{b} + \frac{z_1^2}{c} \leq 1 \\ \frac{x_1^2}{a} + \frac{(1 - y_1)^2}{b} + \frac{z_1^2}{c} \leq 1 \\ \frac{x_1^2}{a} + \frac{y_1^2}{b} + \frac{(1 - z_1)^2}{c} \leq 1 \\ \frac{(1 - x_1)^2}{a} + \frac{(1 - y_1)^2}{b} + \frac{(1 - z_1)^2}{c} \leq 1 \\ \frac{x_1^2}{a} + \frac{y_1^2}{b} + \frac{z_1^2}{c} \geq 1 \\ \frac{(1 - x_1)^2}{a} + \frac{(1 - y_1)^2}{b} + \frac{z_1^2}{c} \geq 1 \\ \frac{x_1^2}{a} + \frac{(1 - y_1)^2}{b} + \frac{(1 - z_1)^2}{c} \geq 1 \\ \frac{(1 - x_1)^2}{a} + \frac{y_1^2}{b} + \frac{(1 - z_1)^2}{c} \geq 1 \end{array} \right. \quad (2)$$

Проте, в загальному випадку система (2) не має розв'язку, тому замість еліпсоїда ми вирішили розглянути квадратичну поверхню, яку можна задати поліномом другого степеня:

$$f(\mathbf{x}) = Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz + J = 0 \quad (3)$$

або в більш компактному вигляді:

$$f(\mathbf{x}) = \mathbf{x}^T M \mathbf{x} + \mathbf{w}^T \mathbf{x} - \gamma, \quad (4)$$

де $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, $M = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix}$.

Розглянемо дихотомію (A, B) . Такий поділ будемо вважати квадратично роздільним, якщо існує квадратична функція $f(\mathbf{x})$ така, що:

$$\begin{aligned} f(\mathbf{a}) &> 0 \quad \forall \mathbf{a} \in A \\ f(\mathbf{b}) &< 0 \quad \forall \mathbf{b} \in B. \end{aligned} \quad (5)$$

Множина точок S_f така, що $S_f = \{\mathbf{x}: f(\mathbf{x}) = 0\}$ – це квадратична поверхня, яка відтинає множину A від множини B [3]. Геометрична задача розділення вершин куба має безпосередній зв'язок з однією з фундаментальних концепцій в комп'ютерних науках, а саме – аналізом булевих функцій. В нашій задачі булева функція трьох змінних – це відображення $g: \{0,1\}^3 \rightarrow \{0,1\}$, тобто існує взаємно однозначна відповідність між 256 комбінаціями вершин куба та 256 можливими булевими функціями трьох змінних. Для будь-якої дихотомії (A, B) ми можемо визначити відповідну булеву функцію g таку, що $g(\mathbf{v}) = 1 \quad \forall \mathbf{v} \in A$ та $g(\mathbf{v}) = 0 \quad \forall \mathbf{v} \in B$ [4]. Завдяки такому ізоморфізму геометрична проблема квадратичної роздільності перетворюється на алгебраїчну, а саме – яка кількість з 256 булевих функцій трьох змінних є квадратичними пороговими функціями?

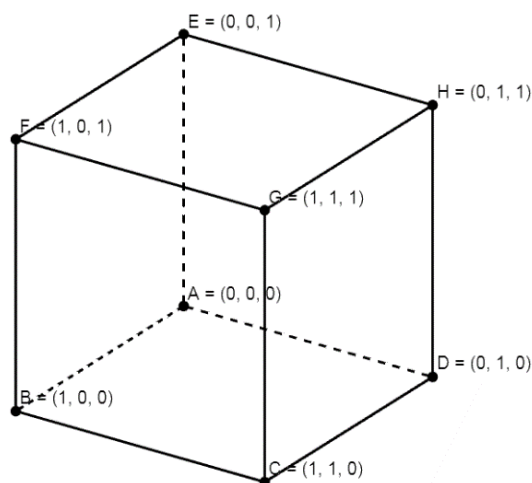


Рис. 2. Модель тривимірного куба Q_3 з вершинами $V = \{0,1\}^3$



3. Ідея застосування теореми Ковера для розв'язання задачі відокремлення множин

В основі нашого підходу лежить фундаментальна ідея з теорії машинного навчання, формалізована теоремою Ковера про підрахунок функцій (Function Counting Theorem), яка дає точну кількість лінійно роздільних дихотомій для N точок у d -вимірному просторі [5]:

$$C(N, d) = 2 \sum_{k=0}^{d-1} \binom{N-1}{k}. \quad (6)$$

Один із найважливіших наслідків цієї теореми, полягає в наступному: задача класифікації, яка не є лінійно роздільною в просторі низької розмірності, з високою ймовірністю стає лінійно роздільною при її нелінійному відображенні у простір вищої розмірності [10].

У тривимірному просторі R^3 загальна квадратична функція має вигляд (3) або (4). Створимо новий простір (простір ознак), в якому кожна координата відповідає одному з доданків нашої функції. Таким чином вектор $\mathbf{x} = (x, y, z)$ трансформується в новий 9-вимірний або 10-вимірний (якщо включати константу) вектор $\phi(\mathbf{x}) = (x^2, y^2, z^2, x, y, z, xy, xz, yz)$. Оскільки для вершин куба V виконується тотожність $x_i^2 = x_i$, то це дозволяє спростити простір, усунувши зайві квадратичні члени. Таким чином, відображення здійснюється в 6-вимірний простір:

$$\phi(x, y, z) \rightarrow (x, y, z, xy, xz, yz). \quad (7)$$

Дихотомія (A, B) є квадратично роздільною у вихідному тривимірному просторі тоді й лише тоді, коли трансформовані множини точок $\phi(A)$ та $\phi(B)$ є лінійно роздільними гіперплощиною в 6-вимірному просторі ознак [6]. Ця концепція є основою так званого «kernel trick» (трюку ядра), що використовується в методі опорних векторів та інших алгоритмах машинного навчання.

Варто зауважити, що теорема Ковера та пов'язані з нею кількісні результати, такі як теорема про підрахунок функцій, спираються на припущення, що точки знаходяться в загальному положенні, тобто жодна їхня підмножина не демонструє «виродженості» (колінеарність або компланарність), що обмежують розділювальні поверхні. Вершини куба – це високоструктурована, симетрична множина точок, і вони не знаходяться в загальному положенні. Через таку структурованість вершин теорема Ковера не

застосовується безпосередньо. Роздільність вершин куба має визначатися конкретним геометричним та комбінаторним аналізом його унікальної структури [7].

Без врахування симетрії існує 256 різних дихотомій куба. Перевіряти їх всі – громіздка задача, але, оскільки куб є об'єктом з високою симетрією, то він володіє такою властивістю, як автоморфізм. Існує 48 різних рухів (поворотів та відображень), які залишають куб незмінним. Ця сукупність рухів утворює групу симетрії куба. За допомогою теореми Пойя [8] встановлено, що серед 256 дихотомій існує 22 унікальні орбіти (геометричні конфігурації вершин куба) (табл. 2).

Таблиця 2

ID орбіти	Розмір	Канонічна множина вершин A (двійкові координати)	Геометричний опис
1	1 vs 7	{(0,0,0)}	Одна вершина
2	2 vs 6	{(0,0,0), (1,0,0)}	Дві суміжні вершини (ребро)
3	2 vs 6	{(0,0,0), (1,1,0)}	Дві вершини на діагоналі грані
4	2 vs 6	{(0,0,0), (1,1,1)}	Дві протилежні вершини
5	3 vs 5	{(0,0,0), (1,0,0), (1,1,0)}	Шлях довжиною 2
6	3 vs 5	{(0,0,0), (1,1,0), (0,1,1)}	Трикутник на діагоналях граней
7	3 vs 5	{(1,0,0), (0,1,0), (0,0,1)}	Рівносторонній трикутник
8	3 vs 5	{(0,0,0), (1,0,0), (0,1,0)}	Три вершини однієї грані (L-форма)
9	4 vs 4	{(0,0,0), (1,0,0), (0,1,0), (1,1,0)}	Грань (квадрат)
10	4 vs 4	{(0,0,0), (1,0,0), (0,1,1), (1,1,1)}	Прямокутник (переріз)
11	4 vs 4	{(1,0,0), (0,1,0), (1,1,1), (0,0,1)}	Ромб
12	4 vs 4	{(0,0,0), (1,1,0), (1,0,1), (0,1,1)}	Правильний тетраедр
13	4 vs 4	{(0,0,0), (0,1,1), (1,0,1), (1,1,0)}	Множина вершин з парною вагою (XOR)
14	4 vs 4	{(0,0,0), (1,0,0), (0,1,0), (1,0,1)}	Три вершини грані та одна суміжна
15	4 vs 4	{(0,0,0), (1,0,0), (1,1,0), (1,1,1)}	Шлях довжиною 3



продовження табл. 2

16	4 vs 4	$\{(1,0,0), (0,1,0), (0,0,1), (1,1,1)\}$	Точки на діагоналях протилежних граней
17	4 vs 4	$\{(0,0,0), (1,0,0), (0,1,1), (1,1,0)\}$	Три точки в площині, одна – ні
18	4 vs 4	$\{(0,0,0), (1,0,0), (0,1,1), (0,1,0)\}$	Дві незв'язані пари суміжних вершин
19	4 vs 4	$\{(0,0,0), (0,1,0), (1,1,1), (1,0,1)\}$	«Скошений» чотирикутник
20	4 vs 4	$\{(0,0,0), (1,0,0), (0,1,0), (0,0,1)\}$	Чотири вершини суміжні з центром
21	4 vs 4	$\{(0,0,0), (1,0,0), (1,1,0), (0,0,1)\}$	«Клин»
22	4 vs 4	$\{(0,0,0), (1,0,0), (1,0,1), (0,1,1)\}$	«Зигзаг»

Якщо одна дихотомія в орбіті є квадратично роздільною, то всі інші дихотомії в цій орбіті також є роздільними. Це означає, що нам потрібно перевірити лише по одному «представнику» з кожної з 22 орбіт.

Трансформуємо координати вершин куба Q_3 в 6-вимірний простір (табл. 3).

Таблиця 3

Вершина	x	y	z	xy	xz	yz
$v_0 = A(0,0,0)$	0	0	0	0	0	0
$v_1 = B(1,0,0)$	1	0	0	0	0	0
$v_2 = C(1,1,0)$	1	1	0	1	0	0
$v_3 = D(0,1,0)$	0	1	0	0	0	0
$v_4 = E(0,0,1)$	0	0	1	0	0	0
$v_5 = F(1,0,1)$	1	0	1	0	1	0
$v_6 = G(1,1,1)$	1	1	1	1	1	1
$v_7 = H(0,1,1)$	0	1	1	0	0	1

Тепер питання роздільності зводиться до розв'язування такої системи нерівностей (5):

$$\begin{cases} \exists \mathbf{w} \in R^6, \exists t \in R^1: \\ \mathbf{w}\mathbf{a} - t \geq 1 \quad \forall \mathbf{a} \in A \\ \mathbf{w}\mathbf{b} - t < 1 \quad \forall \mathbf{b} \in B' \end{cases} \quad (8)$$

де $\mathbf{w} = (w_1, w_2, w_3, w_4, w_5, w_6)$ – вектор ваг (коефіцієнтів шуканого полінома), t – поріг для трансформованих векторів $\phi(\mathbf{x}) = (x, y, z, xy, xz, yz)$. Це задача, що розв'язується в лінійному

програмуванні, яку можна розв'язати автоматично, використовуючи модуль «Розв'язувач» MS Excel. Якщо допустимий розв'язок існує, то дихотомія $\{A, B\}$ є квадратично роздільною, в протилежному випадку така дихотомія не розділяється поліномом другого порядку.

Аналіз показує, що 21 з 22 дихотомій є квадратично роздільними і лише одна орбіта, під номером 16 (табл. 2), яка містить 2 дихотомії, є нероздільною. Таким чином загальна кількість роздільних дихотомій $256 - 2 = 254$. Розділювальні поверхні, які ми знайшли, є прикладами поліноміальних порогових функцій (ППФ) [9]. Наш аналіз показує, які та скільки дихотомій на Q_3 можуть бути представлені ППФ другого ступеня 2 (квадратично роздільні) (табл. 4).

На рис. 3 показано візуальне відображення розділення однієї з дихотомій (орбіта 2) поліномом другого ступеня із знайденими за допомогою розв'язувача коефіцієнтами, а саме: $f(x) = 2yz - 2y - 2z + 1 = 0$.

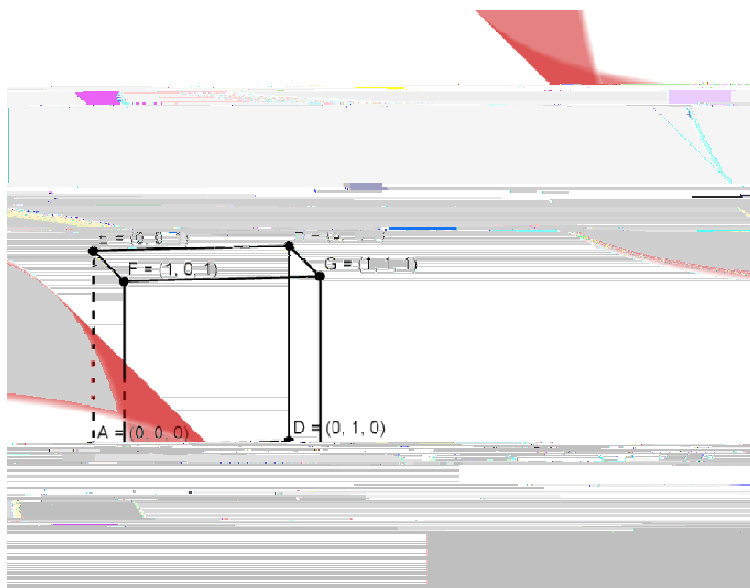


Рис. 3. Візуалізація розділення вершин тривимірного куба Q_3 поліномом $f(x) = 2yz - 2y - 2z + 1 = 0$



Таблиця 4

Клас роздільності	Кількість дихотомій	Відсоток від загальної кількості	Кількість нееквівалентних класів (орбіт)
Загальна кількість дихотомій	256	100%	22
Лінійно роздільні	104	40,6%	8
Квадратично роздільні	254	99,2%	21
Квадратично нероздільні	2	0,8%	1

Задача розділення множин вершин є канонічною задачею розпізнавання образів. Історично, нездатність лінійних моделей (таких як перцептрон) розділяти певні дихотомії куба (особливо XOR) стала рушійною силою для розробки багат шарових перцептронів та нелінійних функцій активації в нейронних мережах. Наш аналіз підтверджує, що перехід від лінійних до квадратичних функцій забезпечує суттєве збільшення класифікаційної потужності, розділяючи 150 із 152 лінійно нероздільних дихотомій.

Дослідження розділення дихотомій тривимірного куба є низькорівневим випадком. Щоб повністю оцінити ландшафт роздільності, необхідно розглянути асимптотичну поведінку при зростанні кількості вимірів. Таке дослідження було проведено в роботі [9], згідно з яким кількість дихотомій $T(n, d)$ вершин n -куба, які є роздільними поліномом степеня d при $n \rightarrow \infty$ визначається, як:

$$\log_2 T(n, d) = \frac{n^{d+1}}{d!} (1 + o(1)). \quad (9)$$

Для оцінки кількості розділяючих функцій розглянемо відношення:

$$\frac{\log_2 T(n, d)}{\log_2 2^{2^n}} \approx \frac{n^{d+1}/d!}{2^n}. \quad (10)$$

Із (10) видно, що при зростанні n частка розділяючих функцій стрімко зменшується. Здатність квадратичних поверхонь розділяти велику кількість дихотомій є особливістю низьких розмірностей. У вищих розмірностях для полінома фіксованого степеня переважна

більшість дихотомій є нероздільними.

Висновки. В статті показано, що застосування квадратичних поліноміальних порогових функцій (PTF-2) майже повністю розв'язують задачу відокремлення вершин куба.

В статті автори використали ідею нелінійного відображення: кожен вершину куба переведено у 6-вимірний простір координат, в якій квадратична роздільність відповідає лінійній роздільності в просторі ознак (теорема Ковера та трюк ядра (kernel trick)).

З огляду на високу симетрію куба, повний перебір 256 дихотомій зведено до 22 нееквівалентних орбіт за теоремою Пойя та чисельно перевірено по одному представнику від кожної орбіти. Задачу роздільності для кожного представника було сформульовано у вигляді системи лінійних нерівностей у просторі ознак, яка була розв'язана за допомогою надбудови «Розв'язувач» у MS Excel.

Встановлено, що 254 із 256 дихотомій є квадратично роздільними (PTF-2), а дві – ні (дихотомія XOR і її доповнення, що відповідають орбіті 16 (табл. 2)). Подано повний класифікаційний підрахунок для PTF-2 на вершинах тривимірного куба із використанням комбінації групових симетрій.

Отриманий авторами результат дає точну межу розділювальної здатності PTF-2 для задач бінарної класифікації у вимірах низької розмірності і пояснює, чому вже перехід від лінійних до квадратичних моделей кардинально підсилює класифікаційну потужність. Наведені асимптотичні оцінки свідчать, що у вимірах з високою розмірністю для полінома фіксованого ступеня частка роздільних дихотомій стрімко спадає. Це підкреслює складність високорозмірних просторів і надає теоретичну основу для необхідності створення альтернативних моделей нейронних мереж.

1. Бичков О. С. Когнітивні методи кібернетики : навч. посіб. К. : Видавничо-поліграфічний центр «Київський університет», 2006. 127 с.
2. Minsky M. and Papert S. *Perceptrons: An Introduction to Computational Geometry*. MIT Press, Cambridge, Mass., 1969.
3. Falk J. E., Dandurova Y., Yeganova L. Set Separation Problems and Global Optimization. *Nonlinear Anal Theory Methods Appl.* 2001 Aug. Vol. 47(3). P. 1857–1867. doi: 10.1016/S0362-546X(01)00316-9. Epub 2001 Jul 31. PMID: 29503498; PMCID: PMC5830190.
4. Shonkwiler R. Separating the vertices of N-cubes by hyperplanes and its application to artificial neural networks. *IEEE Transactions on Neural Networks*. March 1993. Vol. 4, no. 2. Pp. 343–347. doi: 10.1109/72.207621.
5. Cover T. M. Geometrical and Statistical Properties of Linear Threshold Devices. *Technical Report*. Stanford University, 1964. No. 6107-1.
6. Cover T. M. Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition. *IEEE Transactions on Electronic Computers*. 1965. Vol. EC-14(3). P. 326–334.
7. Harary F., Hayes J. P., Horng-Jyh Wu. A survey of the theory of hypercube



graphs. *Computers & Mathematics with Applications*. 1988. Vol. 15, Issue 4. P. 277–289. ISSN 0898-1221. [https://doi.org/10.1016/0898-1221\(88\)90213-1](https://doi.org/10.1016/0898-1221(88)90213-1) **8.** Pólya G. Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen. *Acta Math.* 1937. Vol. 68. P. 145–254. <https://doi.org/10.1007/BF02546665> **9.** Baldi Pierre & Vershynin Roman. Boolean polynomial threshold functions and random tensors. 2018. 10.48550/arXiv.1803.10868. **10.** Vakalopoulou M., Christodoulidis S., Burgos N., et al. Deep Learning: Basics and Convolutional Neural Networks (CNNs) 2023 Jul 23. In: Colliot O, editor. Machine Learning for Brain Disorders [Internet]. New York, NY : Humana; 2023. Chapter 3. URL: <https://www.ncbi.nlm.nih.gov/books/NBK597497/> doi: 10.1007/978-1-0716-3195-9 (дата звернення: 25.02.2025).

REFERENCES:

1. Bychkov O. S. Kohnityvni metody kibernetiky : navch. posib. K. : Vydavnycho-polihrafichnyi tsentr «Kyivskiy universytet», 2006. 127 s. **2.** Minsky M. and Papert S. Perceptrons: An Introduction to Computational Geometry. MIT Press, Cambridge, Mass., 1969. **3.** Falk J. E., Dandurova Y., Yeganova L. Set Separation Problems and Global Optimization. *Nonlinear Anal Theory Methods Appl.* 2001 Aug. Vol. 47(3). P. 1857–1867. doi: 10.1016/S0362-546X(01)00316-9. Epub 2001 Jul 31. PMID: 29503498; PMCID: PMC5830190. **4.** Shonkwiler R. Separating the vertices of N-cubes by hyperplanes and its application to artificial neural networks. *IEEE Transactions on Neural Networks*. March 1993. Vol. 4, no. 2. Pp. 343–347. doi: 10.1109/72.207621. **5.** Cover T. M. Geometrical and Statistical Properties of Linear Threshold Devices. *Technical Report*. Stanford University, 1964. No. 6107-1. **6.** Cover T. M. Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition. *IEEE Transactions on Electronic Computers*. 1965. Vol. EC-14(3). P. 326–334. **7.** Harary F., Hayes J. P., Horng-Jyh Wu. A survey of the theory of hypercube graphs. *Computers & Mathematics with Applications*. 1988. Vol. 15, Issue 4. P. 277–289. ISSN 0898-1221. [https://doi.org/10.1016/0898-1221\(88\)90213-1](https://doi.org/10.1016/0898-1221(88)90213-1) **8.** Pólya G. Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen. *Acta Math.* 1937. Vol. 68. P. 145–254. <https://doi.org/10.1007/BF02546665> **9.** Baldi Pierre & Vershynin Roman. Boolean polynomial threshold functions and random tensors. 2018. 10.48550/arXiv.1803.10868. **10.** Vakalopoulou M., Christodoulidis S., Burgos N., et al. Deep Learning: Basics and Convolutional Neural Networks (CNNs) 2023 Jul 23. In: Colliot O, editor. Machine Learning for Brain Disorders [Internet]. New York, NY : Humana; 2023. Chapter 3. URL: <https://www.ncbi.nlm.nih.gov/books/NBK597497/> doi: 10.1007/978-1-0716-3195-9 (data zvernennia: 25.02.2025).

Turbal Yu. V. [1; ORCID ID: 0000-0002-5727-5334],
Doctor of Engineering, Professor,
Kubai O. V. [1; ORCID ID: 0000-0002-2005-487X],
Post-graduate Student

¹*National University of Water and Environmental Engineering, Rivne*

COVER'S THEOREM AND THE SET-SEPARATION PROBLEM VIA POLYNOMIAL THRESHOLD FUNCTIONS (PTFS)

Separating subsets of points in Euclidean space is a foundational problem in data analysis and neurocomputing; its solvability crucially depends on the class of admissible decision boundaries. Linear models are sharply limited (for three variables, only 104 out of 256 dichotomies are linearly separable), which motivates the transition to nonlinear approaches.

This article presents a complete combinatorial catalogue of all 256 possible separations of the cube's vertices and precisely quantifies the limits of separability by second-degree polynomials. The authors show that, among the 22 unique equivalence classes (orbits) of dichotomies with respect to the cube's symmetry group, exactly one class is not quadratically separable. This analysis not only provides an exhaustive answer to the question posed, but also illustrates fundamental constraints inherent even to powerful nonlinear models when operating on highly structured data.

To this end, we exploit an isomorphism between the geometric separation problem and the algebraic problem of realizing Boolean functions. Each dichotomy of the cube's vertices corresponds uniquely to a Boolean function of three variables, and the question of quadratic separability becomes equivalent to asking whether the given function is a polynomial threshold function of degree two.

The result yields an exact boundary on the expressive power of quadratic threshold elements for binary classification tasks in low dimensions and explains why moving from linear to quadratic models already leads to a dramatic increase in classification capability.

The authors' result establishes the precise limit of the separating power of degree-two polynomial threshold functions (PTF-2) for binary classification in low-dimensional regimes and explains why the transition from linear to quadratic models already markedly enhances classification capacity. The presented asymptotic estimates indicate that, in high-dimensional settings and for a fixed polynomial degree, the proportion of separable dichotomies declines rapidly. This underscores the intrinsic difficulty of high-dimensional spaces and provides a theoretical rationale for developing alternative neural-network models and architectures.



Keywords: polynomial threshold function; cube dichotomy; Cover's theorem; Pólya's theorem; kernel trick; quadratic separators; neural network; Boolean function.

Отримано: 05 травня 2025 року
Прорецензовано: 03 червня 2025 року
Прийнято до друку: 16 червня 2025 року