

НАВЧАННЯ SOM: МЕТОД “НЕЙРОННОЇ МІГРАЦІЇ”

© Годич О.В., 2004

Запропоновано підхід до покращення результативності методу Кохонена навчання нейромереж типу SOM. Цей новий підхід названо „методом нейронної міграції”, оскільки він базується на зміні позицій нейронів у ґратці за певних умов. Для порівняння успішності роботи методу Кохонена та нейронної міграції було проведено експерименти для трьох популярних наборів даних: Animal Database [2], Iris Database [19] та Servo Database [22]. Порівняння базувалося на трьох критеріях – середньо-квадратичному відхиленні (MSE), топографічній помилці (TE) та PSR [17].

Presents an approach for optimisation of the Kohonen’s learning algorithm for Self-Organising Maps. This new approach is called “Neural migration” as it is based on the modification of neurons’ positions in the lattice if certain conditions are met. The comparative simulations between classical Kohonen’s algorithm and “Neural migration” have been performed using three popular data sets: Animal Database [2], Iris Database [19] and Servo Database [22]. The comparison itself was based on three criteria – mean square error (MSE), topographic error (TE) and PSR [17].

Постановка задачі у загальному вигляді

Нейромережі, що самоорганізуються (Self-Organizing Maps, SOM) широко використовуються для розв’язання великої кількості практичних задач. Серед таких задач є кластерний аналіз (cluster analysis) [12], розпізнавання мови (speech recognition) [5], векторне квантування (vector quantization) [10], адаптивна компенсація (adaptive equalization) [6] та інші.

Як і кожна штучна нейронна мережа, SOM має бути навченою для адекватного вирішення поставленої проблеми. Однією з найбільш поширених реалізацій SOM є SOM Кохонена, яка навчається, використовуючи метод Кохонена [7]. Головною властивістю SOM Кохонена є здатність встановлення відношення (mapping) між даними високої розмірності та топологічними структурами нижчої розмірності. Побудова такого відношення відбувається адаптивно топологічно впорядкованим способом (topologically ordered fashion) [7]. У переважній більшості під топологічною структурою нижчої розмірності розуміють двомірну ґратку нейронів [7], яку ще називають картою ознак [2].

Успішна побудова адекватної карти ознак залежить від успішного вибору параметрів методу Кохонена для навчання SOM. Цими параметрами є функція топологічного сусідства та крок навчання, який ще називають параметром швидкості навчання [2]. Якість навчання критично залежить від ініціалізації вагових векторів нейронів перед початком навчання [14].

Аналіз останніх досліджень

Існує декілька підходів покращення роботи методу Кохонена. В основному всі ці підходи зосереджені на питанні побудови ґратки нейронів. Автори праць [1, 3, 11, 13] обговорюють використання генетичних алгоритмів для побудови ґратки нейронів. У праці [8] пропонується метод побудови ґратки нейронів, в якому розмір і форма визначаються динамічно в процесі навчання. Багат шарові SOM були запропоновані у [4]. Вибір функції топологічного сусідства та кроку навчання у переважній більшості визначається експериментальним шляхом. У праці [9] обговорюється питання залежності якості навчання SOM від вибору функції топологічного сусідства.

Цілі статті

У статті розглянуто питання подолання проблеми ініціалізації вагових векторів нейронів у SOM. Автором розроблено метод „Нейронної міграції” (НМ), який є модифікацією методу Кохонена. Базовою ідеєю, яку покладено в основу методу НМ, є адаптивне перегрупування нейронів SOM відповідно до близькості їх вагових векторів. Процес адаптивного перегрупування нейронів названо *нейронною міграцією*. Більш чітко визначення нейронної міграції обговорюється в частині, присвяченій теоретичним аспектам методу. Структурно метод НМ складається з двох процесів: нейронної міграції та адаптації нейронів, де адаптація нейронів SOM відбувається відповідно до методу Кохонена.

Стаття складається із двох основних частин. У першій частині обговорюються теоретичні засади методу НМ та наводиться алгоритм реалізації методу. Друга частина зосереджена на обговоренні ефективності методу НМ, де описуються використані оцінки для перевірки якості SOM, навчених методом НМ та наводяться три проведені симуляції. Ефективність застосування методу НМ оцінюється порівняно із методом Кохонена.

Основний матеріал

Теорія та алгоритм методу НМ

Нехай пара (u, v) унікально визначає позицію нейрон у просторі $U \times V \equiv R^2$. Тобто, $\exists!(u, v) \in U \times V$. Вважатимемо, що кожен нейрон у просторі $U \times V$ є пронумерований і має унікальний порядковий номер $i = \overline{0, l-1}$, де l – загальна кількість нейронів.

Нейрони $i = \overline{0, l-1}$ формують ґратки, де їх координати (u_i, v_i) визначають вузли ґратки (див. рис. 1).

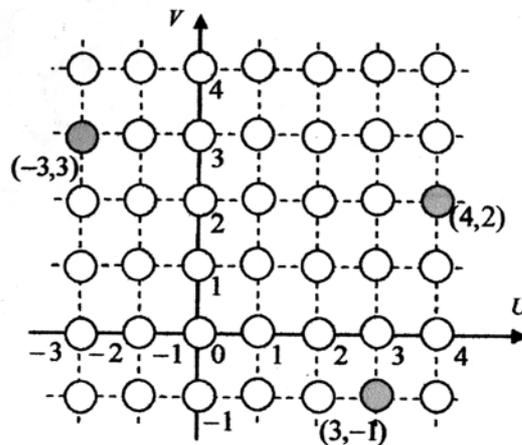


Рис. 1. Схематичне зображення ґратки нейронів у просторі $U \times V$

Нехай топологічне сусідство [2, 7] нейронерезі $\{(u_i, v_i)\}_{i=0}^{l-1}$ задається співвідношенням (1):

$$h_{j,i(x)} = \exp\left(-\frac{d_{j,i(x)}}{2\sigma^2}\right) \quad (1)$$

Співвідношення (1) називається функцією топологічного сусідства і є ідентичним до запропонованого у методі Кохонена [2].

Параметр σ називається ефективною шириною топологічного сусідства, яка визначає кількість нейронів, що увійдуть у топологічне сусідство нейрона-переможця. Величина $d_{j,i(x)}$ – це Евклідова відстань між нейроном-переможцем $i(x)$ та нейроном j :

$$d_{j,i(x)}^2 = \|r_j - r_{i(x)}\|^2, \quad (1^*)$$

де r_j відповідає позиції нейрона j у ґратці, а $r_{i(x)}$ – позиції нейрона $i(x)$. Надалі вхідними даними вважатимемо вектори x з деякого вхідного простору $X \equiv R^n$. Отже, у формулі (1) вектор $x \in X$.

Нейрон-переможець $i(x)$ визначається співвідношенням (2) [2]:

$$i(x) = \arg \min_j \varphi_j(x), \quad (2)$$

де $j = \overline{0, l-1}$. Функція $\varphi_j(x)$ визначається формулою (3):

$$\varphi_j(x) = \|x - w_j\| = \sqrt{\sum_{k=1}^m (x_k - w_{jk})^2}, \quad (3)$$

де $x \in X$, w_j – ваговий вектор нейрона j . У цьому випадку $\varphi_j(x)$ є Евклідовою нормою.

Позначимо топологічне сусідство нейрона-переможця $i(x) = (u^*, v^*)$ при поданому вхідному векторі $x \in X$ як $O(i(x))$. Тобто топологічне сусідство нейрона $i(x)$ будемо розглядати як окіл цього нейрона, який задається співвідношенням (1).

Для метода Кохонена приналежність деякого нейрона ґратки $(u_j, v_j) \in \{(u_i, v_i)\}_{i=0}^{l-1}$, $i = \overline{0, l-1}$ до околу $O(i(x))$ визначається умовою (4*):

$$h_{j,i(x)} > 0. \quad (4^*)$$

Введемо поняття ε -околу нейрона-переможця $i(x)$, яке позначається через $O_\varepsilon(i(x))$ і визначається умовою (4):

$$h_{j,i(x)} > \varepsilon, \quad (4)$$

де $\varepsilon \in [0, 1)$. Величину ε називаємо *порогом топологічного сусідства*. Ця величина використовується для обмеження кількості нейронів, які беруть участь у процесі міграції. При $\varepsilon = 0$ всі нейрони топологічного сусідства будуть брати участь у нейронній міграції. При зростанні ε кількість таких нейронів буде зменшуватись. При $\varepsilon \rightarrow 1$ кількість нейронів, які братимуть участь у процесі міграції, прямує до одиниці – нейрона-переможця. При цьому метод НМ вироджується у метод Кохонена, оскільки нейронна міграція не відбувається.

Під *процесом міграції* розуміємо заміну координат одного нейрона координатами іншого нейрона, причому другий нейрон отримує координати першого (операція *swap*). Зазначимо, що при зміні координат відбувається заміна порядкового номера нейронів. Нейрони, до яких застосовується *swap* операція, утворюють *міграційну пару*.

Нехай маємо довільний $x \in X$ і відповідний йому нейрон-переможець $i(x)$, який є центром околу $O_\varepsilon(i(x))$. І нехай маємо два нейрони з цього околу $j' \in O_\varepsilon(i(x))$ та $j'' \in O_\varepsilon(i(x))$.

Необхідною умовою утворення міграційної пари між нейронами j' та j'' є умова (5):

$$\varphi_{j'}(x) \neq \varphi_{j''}(x). \quad (5)$$

Причому, якщо $\varphi_{j'}(x) < \varphi_{j''}(x)$, то нейрон j' називається домінуючим. І навпаки, якщо $\varphi_{j''}(x) < \varphi_{j'}(x)$, то нейрон j'' є домінуючим. Для забезпечення вищого впорядкування SOM домінуючий нейрон має розташовуватися ближче до нейрона-переможця $i(x)$ у ґратці. Умова (5) вказує, що один із вагових векторів нейронів j' чи j'' є ближчим до вхідного вектора $x \in X$ ніж інший. Вважатимемо нейрон j' домінуючим, тобто виконується умова $\varphi_{j'}(x) < \varphi_{j''}(x)$.

Достатньою умовою утворення міграційної пари між нейронами j' та j'' є умова (6):

$$d_{j',i(x)} \geq d_{j'',i(x)} \quad (6)$$

де $d_{j',i(x)}$ та $d_{j'',i(x)}$ визначається формулою (1*).

Виконання умови (6) вказує на те, що нейрон j'' є ближчим або на однаковій відстані до нейрона-переможця $i(x)$ у ґратці нейронів. Пара нейронів $\{j', j''\} \in O_\varepsilon(i(x))$, для якої виконуються умови (5) та (6), називається *міграційною парою*. Позначимо множину всіх міграційних пар околу $O_\varepsilon(i(x))$ як $M(O_\varepsilon(i(x)))$.

Оскільки нейрон j' є домінуючим, то в ґратці нейронів він має розташовуватись ближче до нейрона-переможця, ніж нейрон j'' . Для досягнення цього застосовується операція $swap(j', j'')$, яка визначається співвідношеннями (7-8):

$$u_{j'} \leftrightarrow u_{j''}, v_{j'} \leftrightarrow v_{j''} \quad (7)$$

$$j' \leftrightarrow j'' \quad (8)$$

Схематично операцію $swap(j', j'')$ наведено на рис. 2. Нейрон-переможець $i(x)$, який подано заштрихованим кружальцем, є центром околу $O_\varepsilon(i(x))$, який, у свою чергу, наведено заштрихованим квадратом. Нейрони з номерами j' та j'' подано суцільно замальованими кружальцями. Лінія зі стрілками по обидва боки, яка з'єднує ці два нейрони, вказує на застосування операції $swap(j', j'')$.

Отже, процес нейронної міграції можна визначити як процес застосування операції $swap(j', j'')$ для $\forall \{j', j''\} \in M(O_\varepsilon(i(x)))$ та $\forall x \in X$.

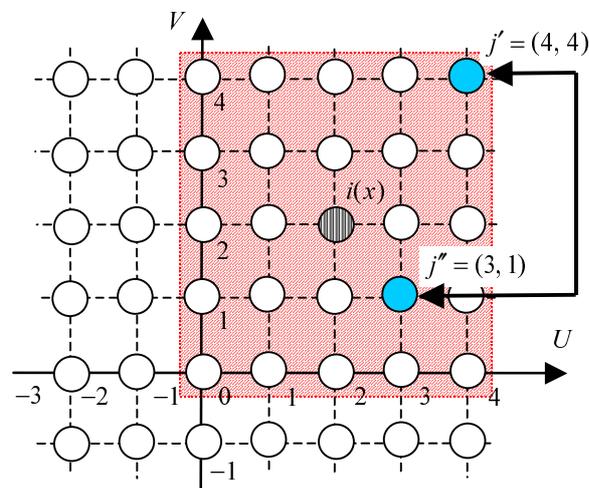


Рис. 2: Схематичне зображення операції $swap(j', j'')$

Щоби визначити, протягом скількох ітерацій методу Кохонена буде застосована нейронна міграція, введемо параметр κ . Застосування нейронної міграції забезпечує покращання навчання SOM лише коли її проводити на етапі організації методу Кохонена [2]. Якщо застосовувати нейронну міграцію протягом етапу збіжності [2], то якість навчання погіршується порівняно із методом Кохонена. Причиною цього є те, що на завершенні етапу організації та протягом всього етапу збіжності остаточно модифікуються вагові вектори нейронів. При цьому топологічні сусідства містять невелику кількість близьких, достатньо добре навчених нейронів. Достатньо добра навченість вживається тут у розумінні того, що до завершення навчання вагові вектори цих нейронів будуть модифіковані незначно порівняно із етапом організації. Якщо в цей час застосовувати нейронну міграцію, то будуть перегруповуватись лише близькі нейрони. Відтак вже сформовані відношення між близькими нейронами порушуються. Загальна рекомендація щодо вибору параметра κ для контролю за тривалістю нейронної міграції є першою половиною етапу організації, коли топологічні сусідства містять велику кількість ненавчених нейронів. Більш точні значення можуть бути отримані експериментальним шляхом.

Отже, застосовуючи нейронну міграцію на κ перших ітераціях етапу організації та модифікуючи вагові нейрони за співвідношеннями метода Кохонена, отримуємо метод „Нейронної міграції”. Повний алгоритм методу НМ, який містить кроки метода Кохонена, подано нижче.

Алгоритм методу НМ можна записати так:

1. *Ініціалізація (initializing)*. Вважатимемо, що компоненти вагових векторів $\{w_j\}_{j=0}^{k-1}$ дорівнюють випадковим числам з рівномірного розподілу на проміжку $[0, 1]$. Задаємо значення порогу топологічно сусідства ε та кількість ітерацій k тривалості нейронної міграції. Покладаємо решту параметрів методу Кохонена відповідно до рекомендацій у праці [2].

2. *Отримання вибірок (sampling)*. Вхідний вектор $x \in X$, який подається на вхід SOM, вибираємо випадково із доступної множини вхідних векторів (можна використати рівномірно розподілену випадкову змінну, яка набуває своїх значень на проміжку натуральних чисел $[1, N]$, де N – кількість векторів у множині, на якій проводиться навчання).

3. *Перевірка подібності (similarity matching)*. Знаходимо нейрон-переможець $i(x)$ для вибраного вхідного вектора $x \in X$ відповідно до формули (2).

4. *Пошук міграційних пар*. Для всіх нейронів $j \in O_\varepsilon(i(x))$ визначаємо міграційні пари, використовуючи умови (5) та (6) і тим формуючи множину міграційних пар $M(O_\varepsilon(i(x)))$.

5. *Міграція нейронів*. Для кожної міграційної пари $(i', i'') \in M(O_\varepsilon(i(x)))$ визначаємо домінуючий нейрон і застосовуємо операцію $swap(j', j'')$, яка визначена співвідношеннями (7) та (8).

6. *Модифікація (update)*. Модифікуємо вагові вектори всіх нейронів у ґратці відповідно до методу Кохонена:

$$w_j(n+1) = w_j(n) + \eta(n)h_{j,i(x)}(n)(x - w_j(n)),$$

де $\eta(n)$ – крок адаптації, n – крок навчання, $h_{j,i(x)}(n)$ – функцією сусідства.

7. *Продовження (continuation)*. Продовжуємо виконання алгоритму з кроку 2, поки не виконається умова зупинки методу Кохонена [2].

Кількісна оцінка методу „Нейронної міграції” (результати симуляції)

Для ілюстрації роботи методу „Нейронної міграції” використано три набори даних (Animal Database, Iris Plants Database, Servo Database), які широко використовуються для перевірки та оцінки методів у кластерному аналізі та розпізнаванні образів. Використані набори даних не є пов'язані між собою і будуть описані за послідовністю їх використання. Результати застосування методу НМ наводяться порівняно із результатами застосування методу Кохонена та ненавченої SOM, яка була використана для навчання. Метод Кохонена для порівняння вибрано тому, що його покладено в основу методу НМ, який, в свою чергу, має покращити роботу методу Кохонена.

При проведенні експериментів на різних наборах даних для забезпечення ідентичних умов роботи методів були використані такі умови:

- Спільна нейромережа, для якої проводилось навчання (для кожного набору даних своя);
- Однакова послідовність подання вхідних векторів впродовж навчання (для кожного набору даних своя).

Зазначимо, що в усіх випадках для навчання використовувалась уся доступна множина вхідних даних.

Спільні параметри для методів Кохонена та НМ, які використовувалися для проведення симуляцій на перелічених наборах даних, наведено у таблиці:

Набір даних	Розмір ґратки	σ_0	b	η_0	τ_2
Animal Database	10×10	4	1000	0.1	10000
Iris Plants Database		7			
Servo Database					

Важливо зазначити, що при такому виборі параметрів топологічне сусідство нейронів зменшується до одного нейрона приблизно на 2000-й ітерації. Отже, етап організації (ordering) триває близько двох тисяч ітерацій. На решті 8000 ітерацій відбувався етап збіжності (convergence).

При обговоренні теоретичних засад методу НМ зверталася увага на потребу застосовувати нейронну міграцію лише на етапі організації. Тривалість нейронної міграції визначається кількістю ітерацій, яка задається параметром k . Значення параметра k визначалось експериментально для кожного використаного набору даних. Компоненти вагових векторів нейронів SOM ініціалізувались випадковими числами з рівномірного розподілу на проміжку $[0, 1]$.

Для оцінки результатів роботи методів НМ та Кохонена використано три числові характеристики:

MSE (*mean square error*) – середньоквадратичне відхилення SOM від вхідного простору [8]. Значення MSE визначає якість апроксимації множини векторів $\chi \subset X$ нейромережею SOM і обчислюється за формулою (19):

$$MSE = \frac{1}{N} \sum_{x \in \chi} \|x - w_{BMU}\|^2, \quad (9)$$

де w_{BMU} – ваговий вектор нейрона-переможця, який є нейроном найкращого наближення (BMU, best matching unit) при поточному значенні вхідного вектора $x \in \chi \subset X$, N – кількість векторів у множині $\chi \subset X$.

TE (*topology preservation*) – збереження топології вхідного простору. Під збереженням топології розуміють властивість SOM відображати близькі вектори близькими нейронами. Факт наведення близьких векторів віддаленими нейронами у SOM називають топологічним збуренням. При збільшенні топологічних збурень ефективність SOM адекватно візуалізувати дані високої розмірності на двовимірній карті (feature map) зменшується. Розглядають декілька способів обчислення TE [15, 16]. У наведеній роботі використовувалась формула (10):

$$TE = \frac{1}{N} \sum_{x \in \chi} \begin{cases} 1, & \|r_{BMU} - r_{SBMU}\| > 1 \\ 0, & \|r_{BMU} - r_{SBMU}\| \leq 1 \end{cases}, \quad (10)$$

де r_{BMU} – координати нейрона найкращого наближення у ґратці нейронів, r_{SBMU} – координати другого нейрона найкращого наближення (SBMU, second best matching unit) у ґратці нейронів.

PSR (*percentage of successful recognitions*) – відсоток успішності розпізнавання векторів вхідного простору X . Значення PSR обчислюється класифікатором, який був запропонований у праці [17]. Загальна концепція роботи класифікатора базується на факті, що нейрони SOM у більшості випадків (в ідеалі завжди) реагують на вектори з того кластера, на розпізнавання якого вони були навчені. Отже, після навчання можна встановити, на який кластер має реагувати кожен із нейронів. Якщо при перевірці нейрон реагує на вектор із визначеного для нього кластера, то вважатимемо реакцію успішною, в іншому випадку – невдалою. Підраховуючи кількість успішних та неуспішних реакцій для кожного нейрона, можна обчислити відсоток успішності реагування кожного нейрона. Загальна успішність SOM обчислюється як зважена сума успішності всіх її нейронів. Значимо, що поняття кластера вживається тут у сенсі множини близьких векторів [17].

Додатково для кожного набору даних проілюстровано карту ознак (feature map) [2] та 3-D поверхню, яка відображає середньоквадратичне відхилення індивідуальних нейронів у SOM від вхідного простору.

Побудова карти ознак базується на тому самому принципі, що й обчислення PSR , з тією відмінністю, що нейронам присвоюється колір конкретного кластера вхідних векторів, на який нейрони навчені реагувати. Це вимагає попереднього присвоєння кольору кожному кластеру векторів.

При побудові 3-D поверхні середньоквадратичного відхилення SOM для кожного i -го нейрона найкращого наближення обчислюється середньоквадратичне відхилення MSE_{BMU}^i його вагового вектора від вхідних векторів, на які він відреагував:

$$MSE_{BMU}^i = \frac{1}{n} \sum_{x \in \chi_i} \|x - w_{BMU}^i\|^2 \quad (11)$$

де w_{BMU}^i – ваговий вектор нейрона найкращого наближення, який реагує на вектори $x \in \chi_i$ деякої підмножини $\chi_i \subset \chi$, n – кількість векторів у підмножині χ_i . Отже, MSE_{BMU}^i є значенням по осі OZ для BMU з номером i , а його координати у ґратці визначають значення по осях OX та OY .

Набір даних Animal Database був вперше запропонований і використаний Ritter та Kohonen у [18]. Він складається із 16 векторів розмірності 13, які утворюють три кластери: трав'яні, хижаки та птахи. Причому атрибут приналежності векторів до конкретного кластера відсутній. Компоненти векторів можуть набувати лише булевих значень.

У результаті проведених експериментів було встановлено оптимальні значення параметрів для методу НМ:

ϵ	K
0.1	600

Тобто, нейронна міграція була проведена на перших 600 ітераціях етапу організації для нейронів, для яких функція топологічного сусідства набувала значення більшого ніж 0.1. Результат навчання SOM методом Кохонена та методом НМ наведено у табл. 1

Таблиця 1

Результат навчання SOM для набору даних Animal Database

Метод	PSR(%)	MSE	TE
НМ	87.5	0.0208217	0
Кохонен	81.3	0.0229379	0
Ненавчена SOM	43.8	1.1342281	0.9375

У табл. 1 також наведені значення оцінок для ненавченої SOM, яка була використана для навчання методами Кохонена та НМ. Гістограма ілюстрація результатів із табл. 1 зображена на рис. 3.

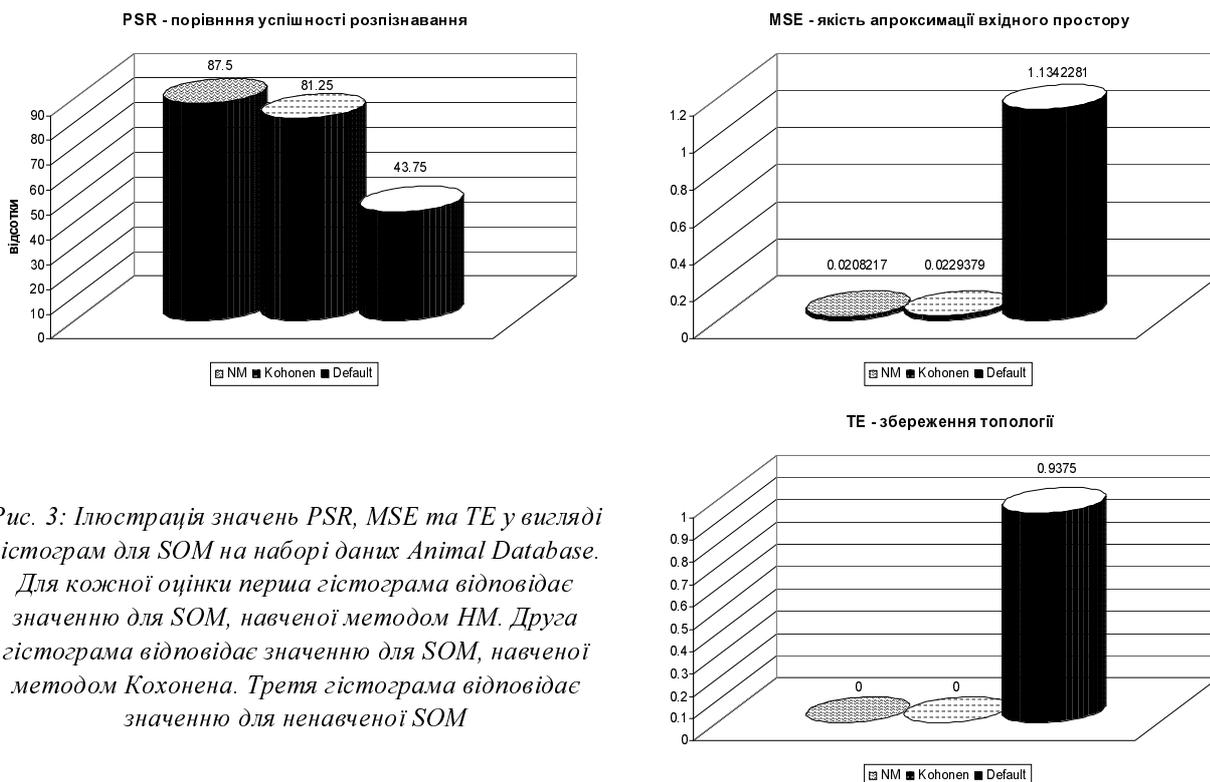


Рис. 3: Ілюстрація значень PSR, MSE та TE у вигляді гістограм для SOM на наборі даних Animal Database.

Для кожної оцінки перша гістограма відповідає значенню для SOM, навченої методом НМ. Друга гістограма відповідає значенню для SOM, навченої методом Кохонена. Третя гістограма відповідає значенню для ненавченої SOM

Як видно із табл. 1 та рис. 3, метод НМ забезпечує кращу апроксимацію вхідного простору і на 6.25% підвищує успішність розпізнавання вхідних векторів. Топологія вхідного простору зберігається однаково SOM навченої методом Кохонена та методом НМ.

На рис. 4 наведено карти ознак для ненавченої SOM, для SOM, навченої методом Кохонена та SOM, навченої методом НМ.

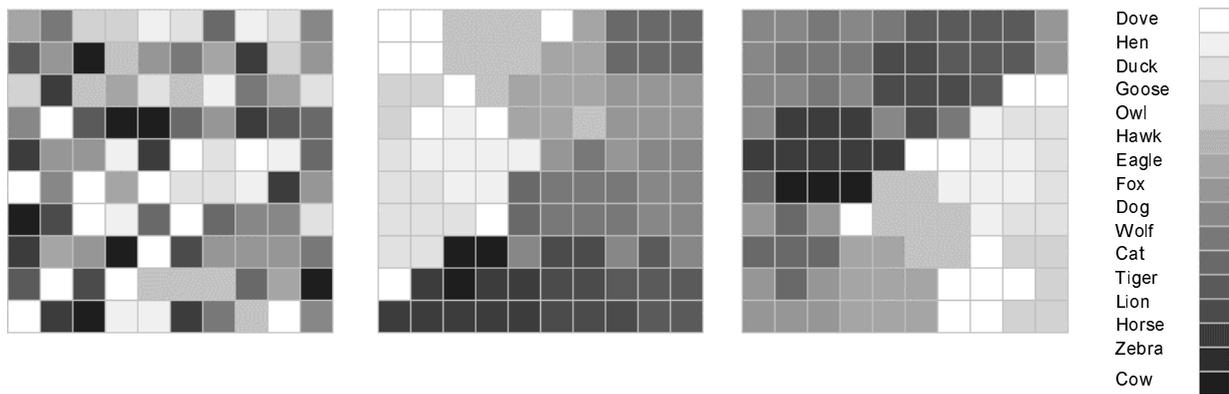


Рис. 4. Зображення карт ознак для SOM до навчання та після навчання на наборі даних *Animal Database* (ненавчена SOM, після навчання методом Кохонена, після навчання методом НМ, легенда кольорів)

Перша карта ознак, зображена на рис. 4 ілюструє SOM до навчання. Візуально легко переконалися, що нейрони, які реагують на травоядних, хижих тварин та птахів, є невпорядковані. Це свідчить про високий рівень топологічного збурення у нейромережі. Після навчання методами Кохонена та НМ карти ознак стабілізуються, що видно з другої та третьої карт ознак на рис. 4. Відмінність між картами ознак для SOM, навчених методом Кохонена та методом НМ, є очевидною, але визначити, яка з карт має нижчий рівень топологічного збурення, неможливо. Для такої оцінки потрібно використовувати критерій ТЕ, який в обох випадках є 0 (див. табл. 1 та рис. 3). Це свідчить про абсолютне відображення топології вхідних даних навченою нейромережею, тобто всі сусідні вхідні вектори наведені сусідніми нейронами.

Основною метою використання карт ознак є проведення візуального аналізу. Зокрема, можна швидко визначити, які дані подані близькими нейронами у ґратці нейронів, а відтак належать до одного кластера.

Однак добра топологічна організованість SOM не свідчить про якісну апроксимацію вхідного простору. І навпаки, добра апроксимація вхідного простору не свідчить про добре топологічне впорядкування нейронів. Якщо карти ознак та критерій ТЕ дозволяють визначати якість SOM у перспективі топологічного впорядкування, то критерій MSE дає можливість оцінити, наскільки добре SOM апроксимує вхідний простір. Додатково пропонується використовувати зображення 3-D поверхні відхилення SOM від вхідного простору. Це забезпечує можливість візуального спостереження за нейромережею в перспективі якості апроксимації вхідного простору. Рис. 5–7 ілюструють такі поверхні для ненавченої SOM та двох SOM, навчених методом Кохонена та НМ відповідно.

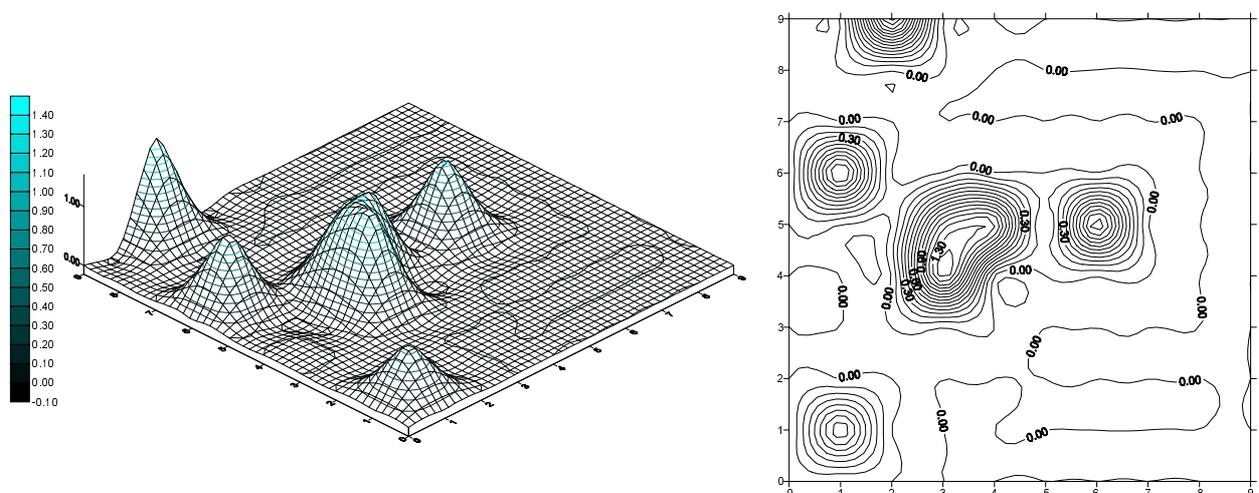


Рис. 5. 3-D поверхня середньоквадратичного відхилення ненавченої SOM для набору даних *Animal Database*

Для ненавченої SOM похибка сконцентрувалась в семи нейронах (див. табл. 2). Слід зазначити, що нейрони, які не відреагували на жоден вхідний вектор мають нульове відхилення.

Таблиця 2

Відхилення нейронів ненавченої SOM на наборі даних Animal Database

Координати нейронів	(1, 1)	(1, 6)	(2, 9)	(3, 4)	(3, 5)	(4, 5)	(6, 5)
Кількість реагувань	3	3	1	1	1	2	5
Середньоквадратичне відхилення	0.77072	1.13041	1.54	1.44795	1.13449	1.22308	1.17513
Квадратичне відхилення (неусереднене)	2.31216	3.39123	1.54	1.44795	1.13449	2.44616	5.87565

Значення середньоквадратичного відхилення ненавченої SOM становить 1.1342281 (див. табл. 11 та рис. 3). Права частина рис. 5 ілюструє лінії рівня, які відповідають побудованій 3-D поверхні.

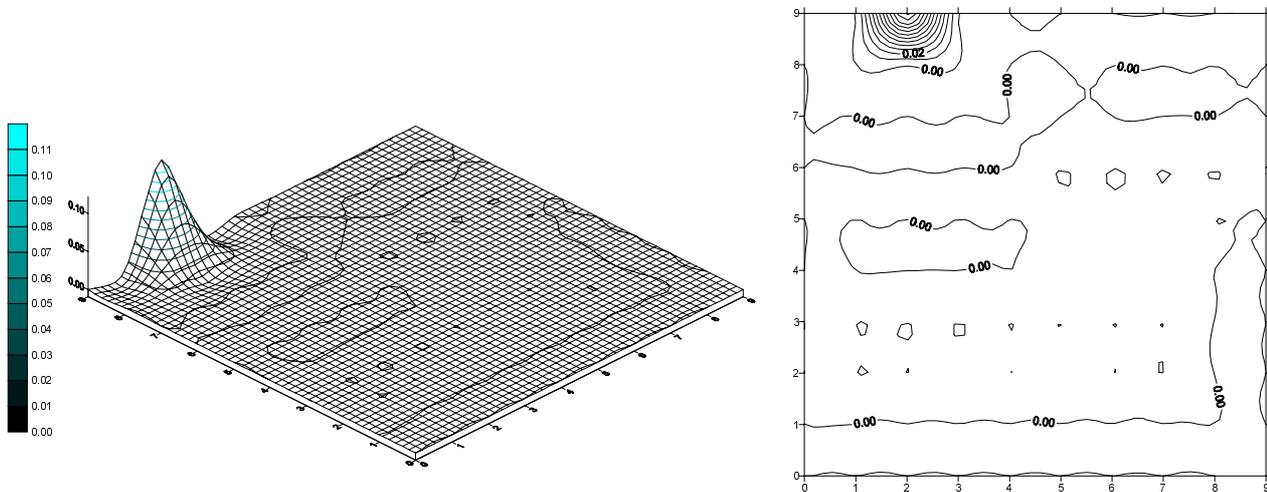


Рис. 6. 3-D поверхня середньоквадратичного відхилення SOM, навченої методом Кохонена для набору даних Animal Database

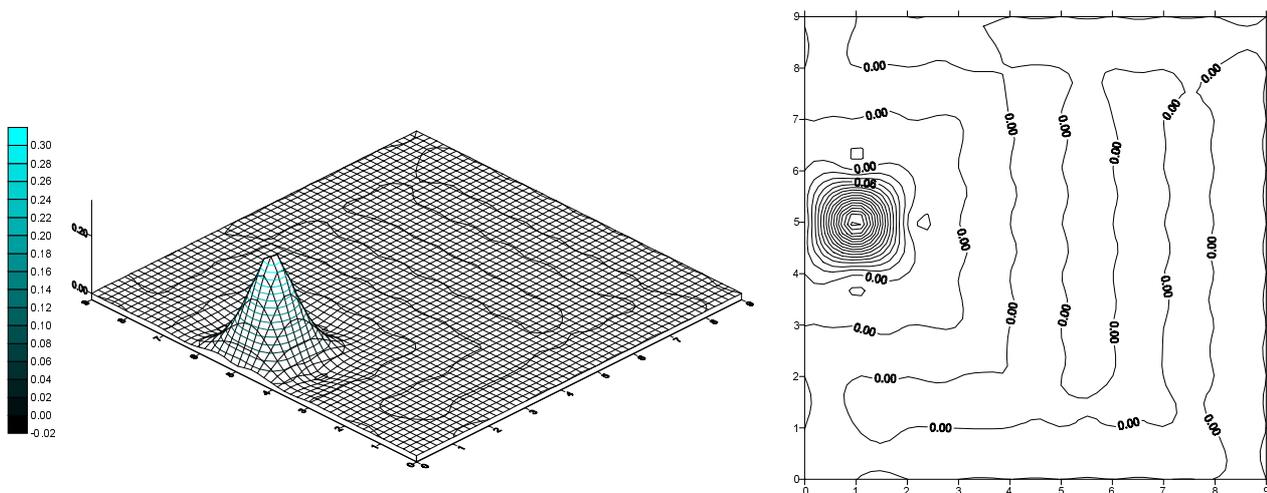


Рис. 7. 3-D поверхня середньоквадратичного відхилення SOM навченої методом HM для набору даних Animal Database

Після навчання методом Кохонена апроксимаційні властивості SOM значно покращилися, і середньоквадратичне відхилення становило 0.0229379 (див. табл. 1 та рис. 3). 3-D поверхня SOM, навченої методом Кохонена та лінії рівня цієї поверхні наведені на рис. 6. Похибка апроксимації накопичилась в одному нейроні, значення якої становить 0.122336 при реагуванні на три вхідні вектори. Тобто неусереднена похибка становить 0.36700683814.

Поверхня середньоквадратичного відхилення SOM, навченої із використанням методу НМ (рис. 7) має схожий вигляд, як для SOM, навченої методом Кохонена (рис. 6). Похибка апроксимації також накопичена в одому нейроні. Значення цієї похибки є вищим і становить 0.333147 при реагуванні на один вхідний вектор. З огляду на реагування цього нейрона лише на один вхідний вектор вплив цієї похибки на загальну результативність є менший, ніж у випадку навчання методом Кохонена. Загальне середньоквадратичне відхилення SOM після навчання методом НМ становило 0.0208217, що вказує на покращення апроксимації вхідного простору порівняно із SOM, навченої методом Кохонена (див. таблицю 1 та рис. 3).

Основною перевагою використання зображення середньоквадратичного відхилення SOM у вигляді 3-D поверхні є легкість візуального виявлення погано навчених нейронів, які є потенційним джерелом нестабільності у навченій SOM. Зокрема, якщо при поданні деякого вектора $x \in \chi$ ВМУ є нейрон з великим значенням похибки порівняно з рештою нейронів, то достовірність такої реакції є сумнівною і вимагає додаткової перевірки.

Набір даних Iris Plants Database є наступним тестовим набором (benchmark), який був використаний для оцінки методу НМ порівняно із методом Кохонена. Цей набір даних був розроблений R.A. Fisher і є одним з найбільш популярних наборів даних у літературі з розпізнавання образів (pattern recognition). Праця Fisher [19] є класичною, і посилання на неї є частими і сьогодні [20, 21].

Набір Iris Plants Database складається із трьох кластерів, кожен з яких містить 50 векторів розмірності 4. Компоненти векторів набувають дійсних значень. Кожен кластер подає окремий тип рослини ірис (iris): Iris-setosa, Iris-versicolor та Iris-virginica. Атрибут приналежності векторів до відповідного кластера є строкового типу і має значення відповідного типу рослини ірис. Кластер Iris-setosa є лінійно роздільний із решта двома (Iris-versicolor та Iris-virginica), які, в свою чергу, є нелінійно роздільні один від одного.

Найкращу результативність навчання SOM методом НМ було досягнуто при значеннях параметрів, які наведено у таблиці:

ϵ	κ
0.3	100

Нейронна міграція проводилась на перших 100 ітерацій етапу організації для нейронів, функція топологічного сусідства яких набувала значення, більшого за 0.3. Результати навчання SOM методом Кохонена, методом НМ та оцінку ненавченої SOM наведено у табл. 3.

Таблиця 3

Результат навчання SOM для набору даних Iris Plants Database

Метод	PSR(%)	MSE	TE
НМ	99.3	0.000094900	0.133333
Кохонен	98.7	0.000111313	0.173333
Ненавчена SOM	68.7	0.053317279	1

На рис. 8 дані з табл. 3 проілюстровано у гістограмному вигляді.

Як можна бачити із табл. 3 та рис. 8, метод НМ забезпечив вищу результативність, ніж метод Кохонена за всіма використаними оцінками. При цьому успішність розпізнавання вхідних векторів для SOM, навчені методом НМ, є на 0.6% вищою.

Карти ознак для ненавченої SOM, SOM, навченої методом Кохонена та SOM, навченої методом НМ, наведено на рис. 9.

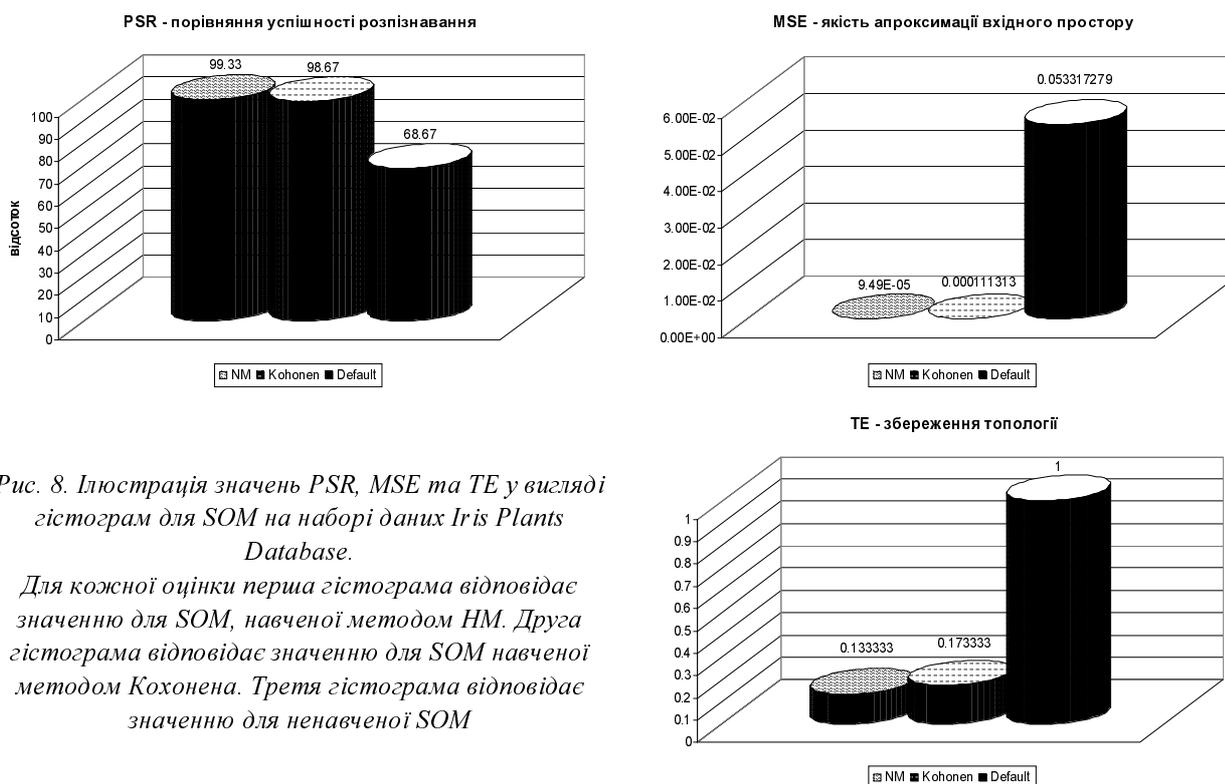


Рис. 8. Ілюстрація значень PSR, MSE та TE у вигляді гістограм для SOM на наборі даних Iris Plants Database.

Для кожної оцінки перша гістограма відповідає значенню для SOM, навченої методом НМ. Друга гістограма відповідає значенню для SOM навченої методом Кохонена. Третя гістограма відповідає значенню для ненавченої SOM

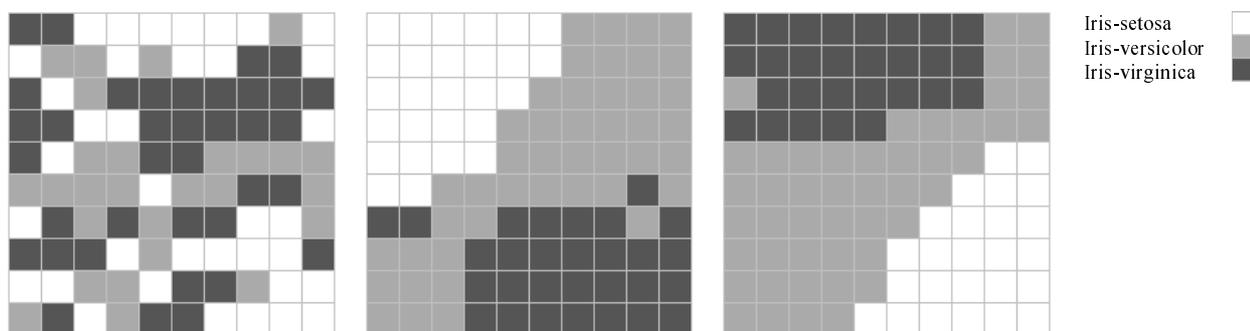


Рис. 9. Зображення карт ознак для SOM до навчання та після навчання на наборі даних Iris Plants Database (ненавчена SOM, після навчання методом Кохонена, після навчання методом НМ, легенда кольорів)

Перша карта ознак на рис. 9 ілюструє ненавчену SOM. Можна бачити, що сусідні нейрони реагують на вектори з різних кластерів, що свідчить про високий рівень топологічного збурення. Значення TE для ненавченої SOM становило 1. Друга та третя карти ознак належать SOM, навченим методом Кохонена та методом НМ відповідно. На цих зображеннях чітко видно лінійно відділений кластер Iris-setosa та два нелінійно розділені кластери Iris-versicolor та Iris-virginica. При цьому SOM, навчена методом НМ, краще подає нелінійно роздільні кластери Iris-versicolor та Iris-virginica, що забезпечило краще збереження топології вхідного простору (див. табл. 3 та рис. 8).

Поверхню середньоквадратичного відхилення та її лінії рівня для ненавченої SOM проілюстровано на рис. 10.

У ненавченій SOM накопичення похибки відбулося у п'яти нейронах. Значення відхилення цих нейронів та кількість векторів, на які вони зреагували, наведено у табл. 4.

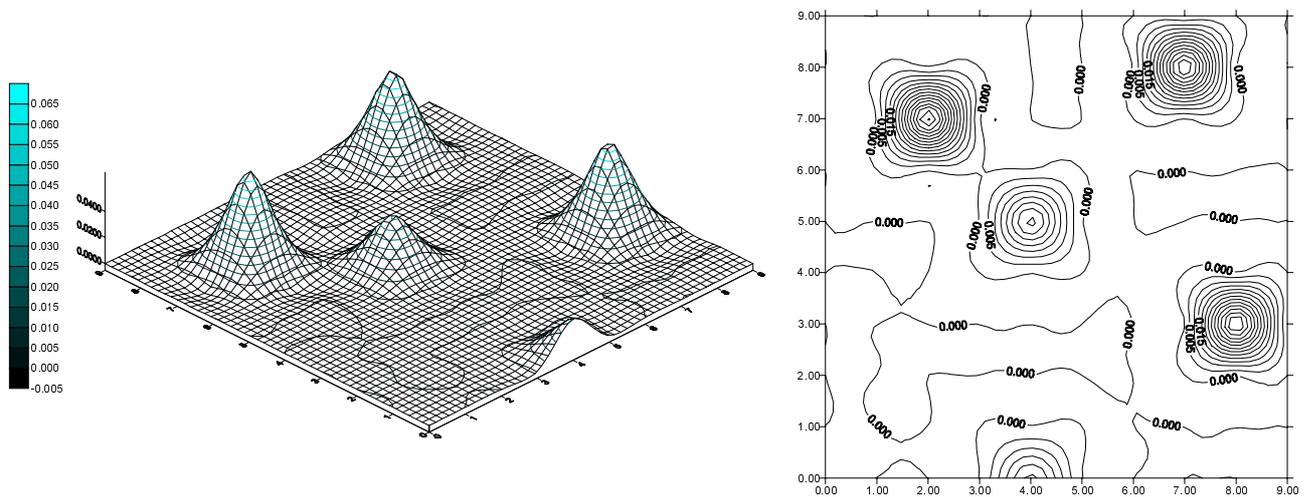


Рис. 10. 3-D поверхня середньоквадратичного відхилення ненавченої SOM для набору даних Iris Plants Database

Таблиця 4

Відхилення нейронів ненавченої SOM на наборі даних Iris Plants Database

Координати нейронів	(2, 7)	(4, 0)	(4, 5)	(7, 8)	(8, 3)
Кількість реагувань	1	38	12	97	2
Середньоквадратичне відхилення	0.070896	0.026484	0.036787	0.065442	0.065474
Квадратичне відхилення (неусереднене)	0.070896	1.006403	0.441446	6.347893	0.130948

Всі решта 95 нейронів, окрім п'яти, наведених у табл. 4, не зреагували жодного разу. Значення MSE для ненавченої SOM становило 0.053317279 (див. табл. 3 та рис. 8).

Для SOM, навченої методом Кохонена, поверхня середньоквадратичного відхилення зображена на рис. 11.

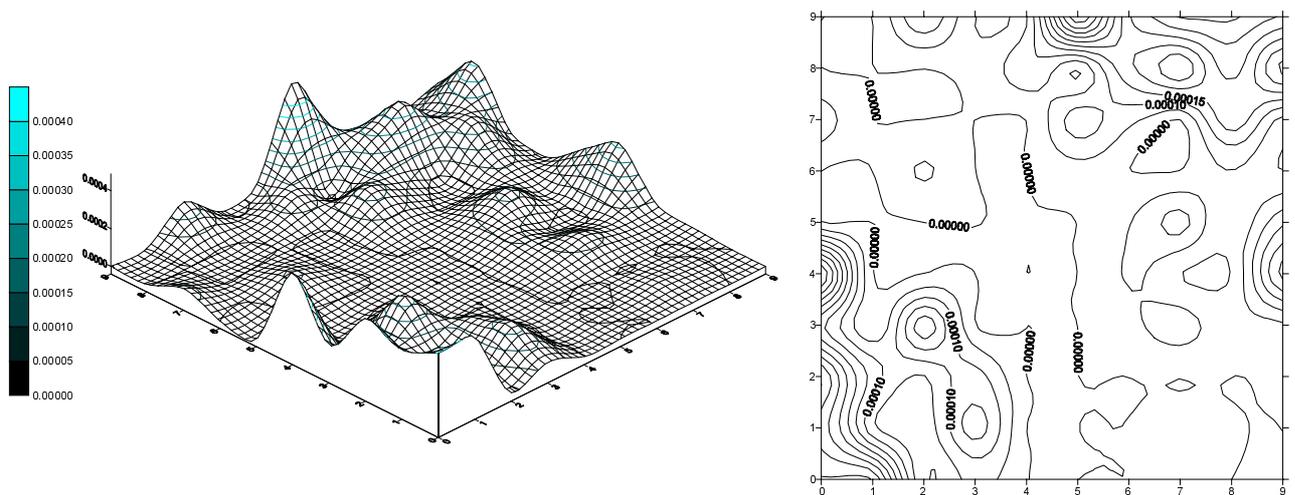


Рис. 11. 3-D поверхня середньоквадратичного відхилення SOM, навченої методом Кохонена для набору даних Iris Plants Database

Для цієї SOM жодного разу не зреагувало 25 нейронів. Накопичення похибки апроксимації вхідного простору відбулось на 23-х нейронах. У табл. 5 наведено перші 8 нейронів, які відповідають найбільшим похибкам. Решта 52 нейрони мали середньоквадратичне відхилення порядку 10^{-5} і нижче.

**Відхилення нейронів SOM навченої методом Кохонена
на наборі даних Iris Plants Database**

Координати нейронів	(7, 8)	(9, 8)	(0, 1)	(1, 0)	(0, 2)	(0, 0)	(0, 4)	(5, 9)
Кількість реагувань	4	3	3	2	2	3	2	2
Середньоквадратичне відхилення	0.000295	0.000313	0.000328	0.000396	0.000398	0.000405	0.000448	0.000495
Квадратичне відхилення (неусереднене)	0.001179	0.000940	0.000983	0.000791	0.000796	0.001214	0.000895	0.000989

Значення MSE для SOM, навченої методом Кохонена, становило 0.000111313 (див. табл. 3 та рис. 8).

На рис. 12 зображена поверхня середньоквадратичного відхилення SOM, навченої методом НМ.

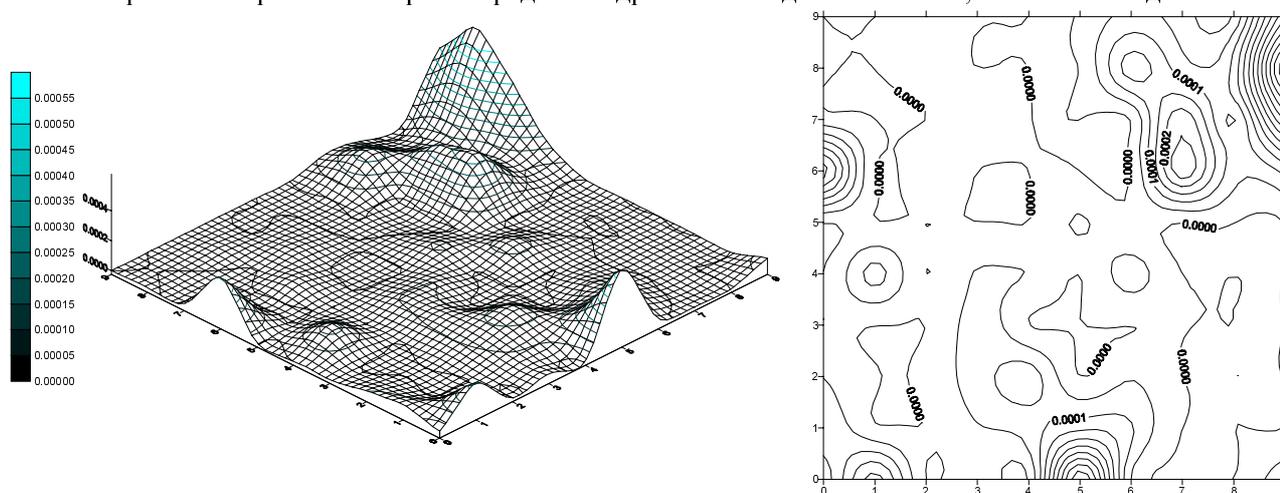


Рис. 12. 3-D поверхня середньоквадратичного відхилення SOM, навченої методом НМ для набору даних Iris Plants Database

Кількість нейронів, які не брали участі у розпізнаванні вхідних векторів, у цьому випадку становила 24. Основне накопичення похибки відбулось на 15-ти нейронах. Для порівняння із результатами методу Кохонена, 8 нейронів з найбільшим відхиленням наведено у табл. 6. Решта 61 нейронів зафіксували середньоквадратичне відхилення порядку 10^{-5} і нижче.

Таблиця 6

**Відхилення нейронів SOM, навченої методом НМ
на наборі даних Iris Plants Database**

Координати нейронів	(1, 0)	(7, 7)	(7, 6)	(0, 6)	(9, 9)	(9, 7)	(5, 0)	(9, 8)
Кількість реагувань	6	3	2	2	2	3	2	2
Середньоквадратичне відхилення	0.000221	0.000235	0.000284	0.000320	0.000396	0.000411	0.000495	0.000653
Квадратичне відхилення (неусереднене)	0.001326	0.001176	0.000567	0.000639	0.000791	0.001233	0.000989	0.001306

Отже, SOM, навчена методом Кохонена, містить більше нейронів з більшим відхиленням, ніж SOM, навчена методом НМ. Значення MSE для SOM, навченої методом НМ, становило $9.49E-05$ (див. табл. 3 та рис. 8).

Набір даних Servo Database був отриманий Karl Ulrich (MIT) у 1986 р. при проведенні симуляції деякої серво-системи. Вперше цей набір даних був використаний J.R. Quinlan у праці [22] у 1992 році. Інше використання цих даних наводилось у праці [23]. J.R. Quinlan описує ці дані як „extremely non-linear”, звертаючи увагу на високу нелінійну роздільність кластерів. Цей набір даних містить 167 векторів розмірності 4. Компоненти векторів набувають свої значення із множини натуральних чисел.

На відміну від попередніх розглянутих наборів даних, Servo Database має числовий атрибут приналежності до того чи іншого кластера. Отже, можна легко використати цей атрибут як п'яту компоненту вхідних векторів. Однак, у проведених експериментах атрибут приналежності використовувався лише для побудови карт ознак, як і для попередніх випадків.

Проведені експерименти визначили оптимальні параметри для методу НМ:

ϵ	K
0.4	500

Відповідно до цих значень процес нейронної міграції відбувався протягом перших 500 ітерацій етапу організації для нейронів, функція топологічного сусідства яких набувала значення, більше за 0.4. Результат навчання SOM методом Кохонена та методом НМ наведено у табл. 7.

Таблиця 7

Результат навчання SOM для набору даних Servo Database

Метод	PSR(%)	MSE	TE
НМ	65.3	0.00227183	0.149701
Кохонен	63.5	0.00220009	0.191617
Ненавчена SOM	29.9	0.04225029	0.940120

Як і для решти використаних наборів даних, значення для кожної із оцінок ілюструються, використовуючи гістограмне подання (рис. 13).

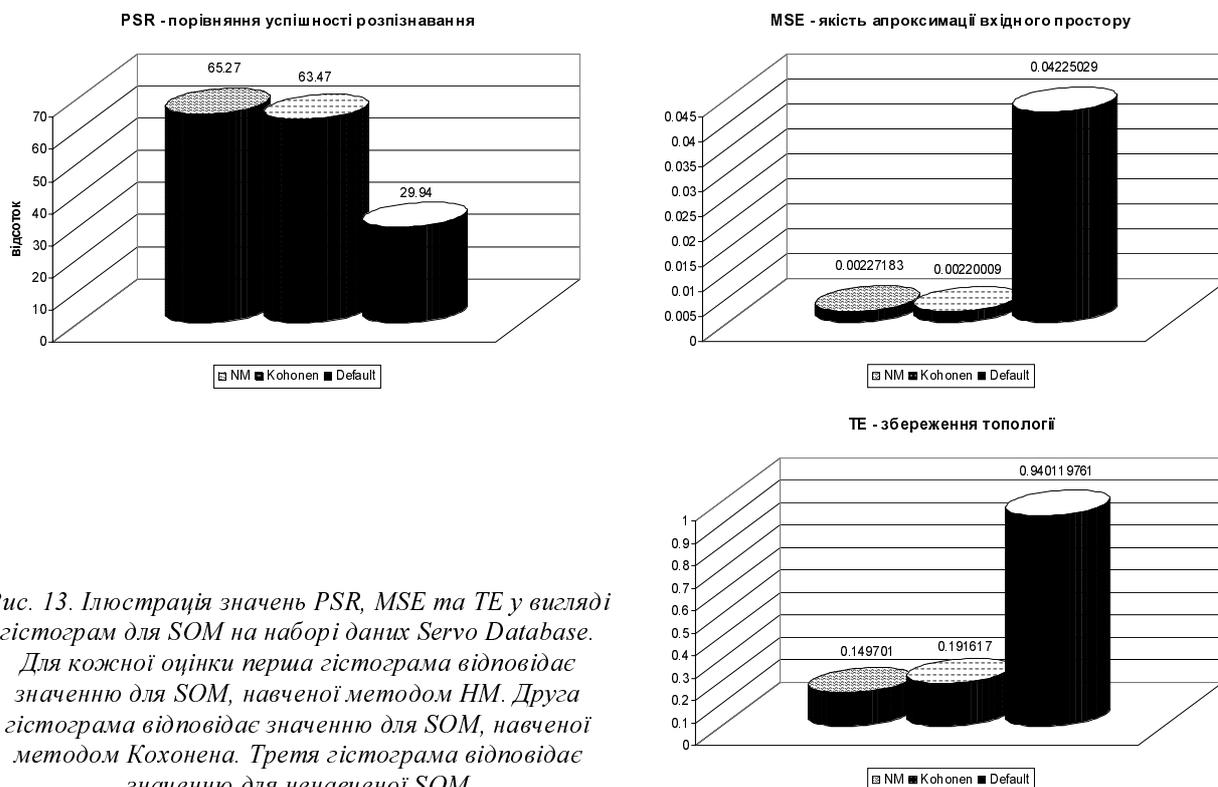


Рис. 13. Ілюстрація значень PSR, MSE та TE у вигляді гістограм для SOM на наборі даних Servo Database. Для кожної оцінки перша гістограма відповідає значенню для SOM, навченої методом НМ. Друга гістограма відповідає значенню для SOM, навченої методом Кохонена. Третя гістограма відповідає значенню для ненавченої SOM

Значення оцінок в табл. 7 та рис. 13 вказують на дещо гіршу апроксимацію вхідного простору для SOM, навчену методом НМ, ніж методом Кохонена. Водночас збереження топології є кращим саме для SOM, навченої методом НМ. Це дало змогу досягнути на 1.8% вищу результативність при розпізнаванні вхідних векторів.

З огляду на велику кількість кластерів (51) карти ознак, наведені на рис. 14, реально не дають можливості візуально оцінити збереження топології для SOM використаного розміру (10×10).

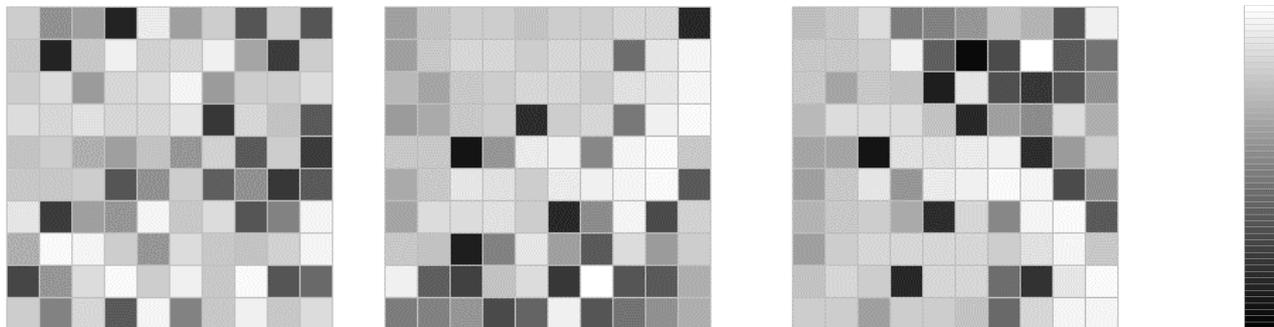


Рис. 14. Зображення карт ознак для SOM до навчання та після навчання на наборі даних Servo Database (ненавчена SOM, після навчання методом Кохонена, після навчання методом НМ, легенда кольорів)

Світлі кольори відповідають меншим значенням, а темні – більшим значенням міток, якими є числа з проміжку [0.13, 7.10]. Зауважимо, що кольори у легенді не підписані через велику кількість міток.

Легко переконатися, що досить важко з впевненістю стверджувати, яка з SOM (навчена методом Кохонена чи НМ) є більш організована. Водночас нижча організованість ненавченої SOM визначається візуально. Значення TE для SOM, навченої методом НМ та Кохонена, становило 0.149701 та 0.191617 відповідно (див. табл. 7 та рис. 13).

Поверхня середньоквадратичного відхилення для ненавченої SOM на наборі даних Servo Database проілюстрована на рис. 15.

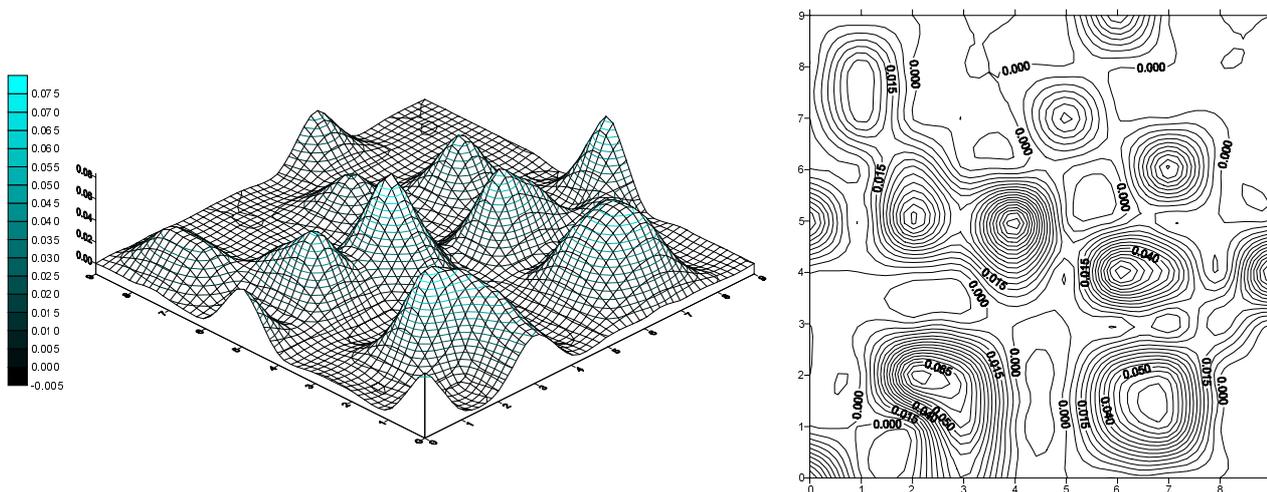


Рис. 15. 3-D поверхня середньоквадратичного відхилення ненавченої SOM для набору даних Servo Database

Накопичення похибки у ненавченій SOM відбулось у 24-х нейронах. Решта 76 нейронів не брали участі у розпізнаванні вхідних векторів. Значення MSE для ненавченої SOM становило 0.04225029 (див. табл. 7 та рис. 13). Значення відхилення перших семи нейронів з найбільшим відхиленням наведено у таблиці 8.

Відхилення нейронів ненавченої SOM на наборі даних Servo Database

Координати нейронів	(7, 1)	(3, 2)	(2, 5)	(6, 4)	(9, 4)	(2, 2)	(4, 5)
Кількість реагувань	1	5	3	2	1	3	4
Середньоквадратичне відхилення	0.059862	0.060728	0.061829	0.068399	0.068719	0.076703	0.084149
Квадратичне відхилення (неусереднене)	0.059862	0.303642	0.185488	0.136798	0.068719	0.230108	0.336596

Для SOM, навченої методом Кохонена на наборі даних Servo Database, поверхня середньоквадратичного відхилення проілюстрована на рис. 16.

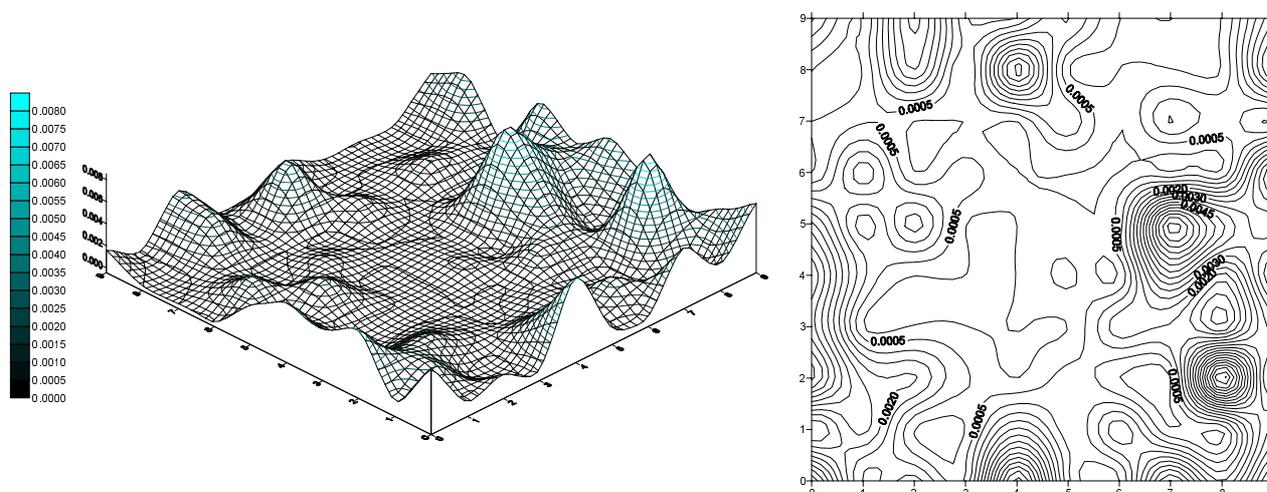


Рис. 16. 3-D поверхня середньоквадратичного відхилення SOM, навченої методом Кохонена на наборі даних Servo Database

У цьому випадку не відреагувало лише 4 нейрони, що свідчить про добрий розподіл реакцій між нейронами. Основна маса похибок сконцентрувалась у 64-х нейронах. Решта 32 нейрони зафіксували середньоквадратичне відхилення порядку 10^{-5} і нижче. Перші 7 нейронів з найбільшим зафіксованим середньоквадратичним відхиленням подані у табл. 9.

Відхилення нейронів SOM навченої методом Кохонена на наборі даних Servo Database

Координати нейронів	(9, 6)	(0, 0)	(0, 2)	(9, 0)	(4, 0)	(7, 5)	(8, 2)
Кількість реагувань	2	4	2	2	2	3	3
Середньоквадратичне відхилення	0.005187	0.00544	0.005644	0.005843	0.007244	0.007688	0.008642
Квадратичне відхилення (неусереднене)	0.010374	0.021762	0.011287	0.011686	0.014488	0.023064	0.025925

Для SOM, навченої методом НМ на наборі даних Servo Database, поверхня середньоквадратичного відхилення проілюстрована на рис. 17.

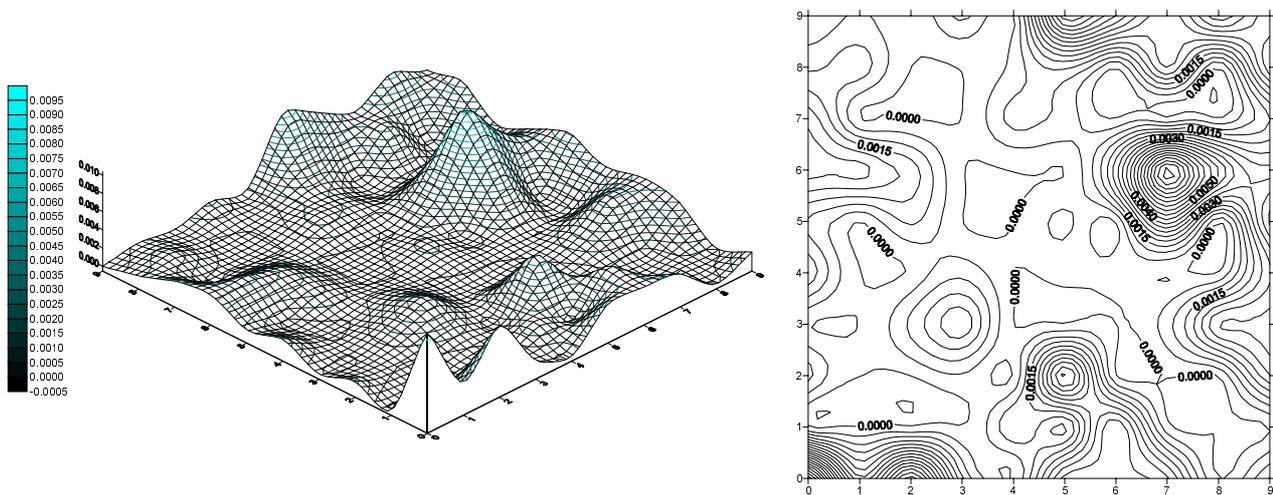


Рис. 17. 3-D поверхня середньоквадратичного відхилення SOM, навченої методом НМ на наборі даних Servo Database

Після навчання методом НМ пасивними залишилися п'ять нейронів – це на один нейрон більше, ніж для методу Кохонена. Основна маса похибок сконцентрувалась у 65 нейронах. Решта 30 нейронів зафіксували середньоквадратичне відхилення порядку 10^{-5} і нижче. Перші 7 нейронів з найбільшим середньоквадратичним відхиленням подано у табл. 10.

Таблиця 10

Відхилення нейронів SOM, навченої методом НМ на наборі даних Servo Database

Координати нейронів	(8, 6)	(5, 2)	(8, 9)	(5, 9)	(2, 0)	(7, 6)	(0, 0)
Кількість реагувань	3	2	3	2	3	4	3
Середньоквадратичне відхилення	0.005027	0.005125	0.005337	0.006832	0.007264	0.009250	0.010377
Квадратичне відхилення (неусереднене)	0.015081	0.010251	0.016010	0.013665	0.021791	0.037002	0.031130

Порівнюючи дані у табл. 9 та 10, можна бачити, що середньоквадратичне відхилення нейронів для SOM, навченої методом НМ, є дещо вище. У результаті значення MSE для цієї SOM є на 0.00007174 більшим.

Висновки

У статті запропоновано підхід до покращання роботи методу Кохонена, в основі якого лежить адаптивне перегрупування нейронів у ґратці впродовж навчання SOM з метою досягнення вищого топологічного впорядкування. Запропонований метод названо методом „Нейронної міграції” (НМ), де під нейронної міграцією розуміємо зміну положень нейронів у ґратці відповідно до розроблених правил.

У розділі “Теорія та алгоритм” детально описуються теоретичні та алгоритмічні аспекти методу НМ. Звертається особлива увага на неможливість застосування нейронної міграції впродовж всього процесу навчання.

Проведені симуляції (simulation) та оцінка (evaluation) запропонованого методу вказує на підвищення класифікаційних можливостей SOM порівняно із класичним методом Кохонена. Важливо зазначити, що як тестові приклади були використані відомі і широко вживані набори даних, які використовуються саме з метою оцінювання методів у кластерному аналізі та розпізнаванні образів. Відмінність у використаних наборах даних (розмірність, тип значень, кількість кластерів тощо) забезпечує різнобічне тестування і дає змогу перевірити застосовність методу у різних випадках.

Табл. 11 містить зведений результат застосування методів НМ та Кохонена для всіх використаних тестових наборів даних.

Зведений результат симуляції для трьох тестових наборів даних

	Animal Database			Iris Plans Database			Serve Database		
	PSR(%)	MSE	TE	PSR(%)	MSE	TE	PSR(%)	MSE	TE
НМ	87.5	0.0208217	0	99.3	0.000094900	0.133333	65.3	0.00227183	0.149701
Кохонен	81.3	0.0229379	0	98.7	0.000111313	0.173333	63.5	0.00220009	0.191617
Різниця (НМ - Кохонен)	6.2	-0.0021162	0	0.6	-0.000016413	-0.04	1.8	0.00007174	-0.041916

Аналізуючи результати, наведені у таблицях та рисунках у розділі „Кількісна оцінка методу „Нейронної міграції” (результати симуляції), а також у табл. 11, можна перекоонатись, що запропонований метод „Нейронної міграції” підвищує класифікаційні властивості SOM та забезпечує краще збереження топології даних в усіх проведених тестах. Для Animal Database та Iris Plants Database метод НМ забезпечує незначне покращання апроксимації вхідних даних. Водночас для Servo Database якість апроксимації незначно погіршилась, що не вплинуло на покращання класифікаційних властивостей SOM.

1. S.A. Harp and T. Samad, "Genetic optimization of self-organizing feature maps," in *Proc. Int. Conf. Neural Networks, Seattle, WA, 1991*, pp. 341–346. 2. Simon Haykin. "Neural Networks: A comprehensive foundation". Prentice-Hall, New Jersey, 1999. 3. S. J. Huang and C. C. Hung, "Genetic algorithms enhanced Kohonen's neural networks," in *Proc. IEEE Int. Conf. Neural Networks, 1995*, pp. 708–712. 4. J. Koh, M. Suk, and S. M. Bhandarkar, "A multilayer self-organizing feature map for range image segmentation," *Neural Networks*, vol. 8, no. 1, pp. 67–86, 1995. 5. T. Kohonen, "The 'neural' phonetic typewriter," *Computer*, vol. 21, pp. 11–22, Mar. 1988. 6. T. K. Kohonen, J. Kangas, J. Laaksonen, and K. Tokkola, "LVQ-PAK: The learning vector quantization program package", *Helsinki Univ. Technol., Finland*, 1992. 7. T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biol. Cybern.*, vol. 43, pp. 59–69, 1982. 8. J. Si, S. Lin, M.-A. Vuong, "Dynamic topology representing networks", *Neural Networks*, vol. 13, 2000, pp.617-627. 9. Z.-P. Lo and B. Bavarian, "On the rate of convergence in topology preserving neural networks," *Biol. Cybern.*, vol. 65, pp. 55–63, 1991. 10. S. P. Luttrell, "Hierarchical vector quantization," *Proc. Inst. Elect. Eng.*, vol. 136, pp. 405–413, 1989. 11. M. McInerney and A. Dhawan, "Training the self-organizing feature map using hybrids of genetic and Kohonen methods," in *Proc. IEEE Int. Conf. Neural Networks, 1994*, pp. 641–644. 12. M. C. Su, N. DeClaris, and T. K. Liu, "Application of neural networks in cluster analysis," in *Proc. IEEE Int. Conf. Systems, Man, and Cybernetics, Orlando, FL, Oct. 12–15, 1997*, pp. 1–6. 13. M. C. Su and H. T. Chang, "Genetic-algorithm-based approach to self organizing feature map and its application in cluster analysis," in *Proc. IEEE Int. Joint Conf. Neural Networks, AK, 1995*, pp. 2116–2121. 14. Mu-Chun Su and Hsiao-Te Chang, "Fast Self-Organizing Feature Map Algorithm", *IEEE Transactions on neural networks*, vol. 11, No. 3, May 2000. 15. T. Villmann, H. Bauer, M. Herrmann, "Neural Maps and Topographic Vector Quantization", *Neural Networks*, vol. 12, pp. 659-676, 1999. 16. T. Villiman, M. Herrmann, R. Der and M. Martinetz, "Topology Preservation in Self-Organizing Feature Map: Exact Definition and Measurement", *IEEE Transactions on Neural Networks*, vol. 8, no. 2, MAR 1997. 17. Годич О., Нікольський Ю., Щербина Ю. Застосування штучної нейронної мережі типу SOM для розв'язування задачі діагностування // Вісник національного університету "Львівська політехніка". – 2002. – № 464. – С. 31–43. 18. Ritter, H., and T. Kohonen, "Self-organizing semantic maps", *Biological Cybernetics*, vol. 61, pp. 241–254. 19. Fisher, R.A. "The use of multiple measurements in taxonomic problems", *Annual Eugenics*, 7, Part II, 179-188 (1936); also in "Contributions to Mathematical Statistics" (John Wiley, NY, 1950). 10. Duda, R.O., & Hart, P.E. (1973) *Pattern Classification and Scene Analysis*. (Q327.D83) John Wiley & Sons. ISBN 0-471-22361-1. See page 218. 21. Dasarathy, B.V. (1980) "Nosing Around the Neighborhood: A New System Structure and Classification Rule for Recognition in Partially Exposed Environments". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-2, No. 1, 67–71. 22. Quinlan, J.R., "Learning with continuous classes", *Proc. 5th Australian Joint Conference on AI* (eds A. Adams and L. Sterling), Singapore: World Scientific, 1992. 23. Quinlan, J.R., "Combining instance-based and model-based learning", *Proc. ML'93* (ed P.E. Utgoff), San Mateo: Morgan Kaufmann 1993.