

ДОДАТОК А

Графічний матеріал кваліфікаційної роботи

Харківський національний університет радіоелектроніки

Кваліфікаційна робота

Метод кластеризації даних з використанням машинного навчання

Виконав:
студент гр. СПзм-20-1
Хануков П.Д.

Керівник:
доц. каф. ЕОМ
Ткачов В.М.

Мета та завдання роботи

2

Метою кваліфікаційної роботи є дослідження ефективності кластеризації даних за допомогою методів машинного навчання, зокрема карт Кохонена.

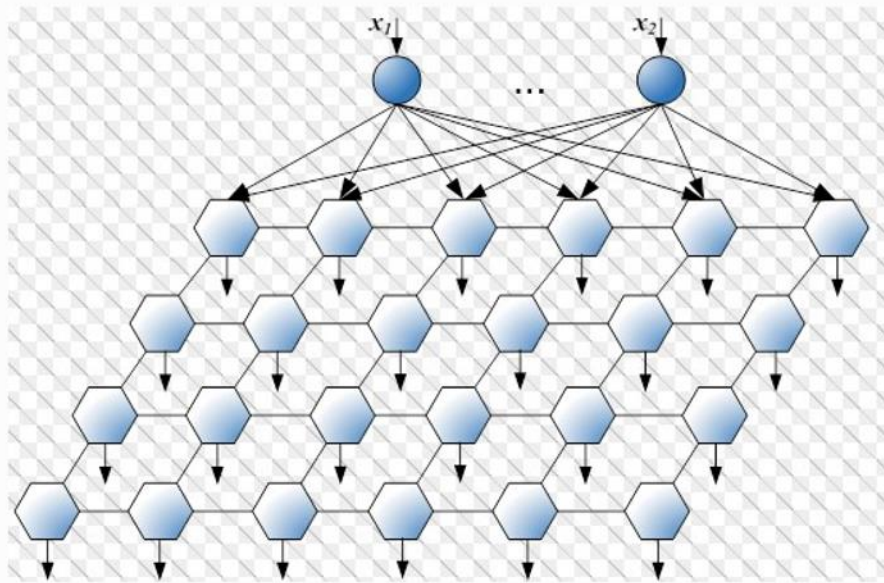
Об'єктом дослідження є карти Кохонена з модифікованою структурою.

Завдання:

- дослідження методів кластеризації даних;
- дослідження методів машинного навчання;
- дослідження алгоритму роботи класичної моделі нейронної мережі Кохонена, його недоліків і існуючих способів їх усунення;
- розробка методу удосконалення апарату нейронної мережі Кохонена для усунення недоліків мереж цього типу: граничного ефекту і появи "мертвих" нейронів;
- проведення моделювання з використанням розробленого методу.

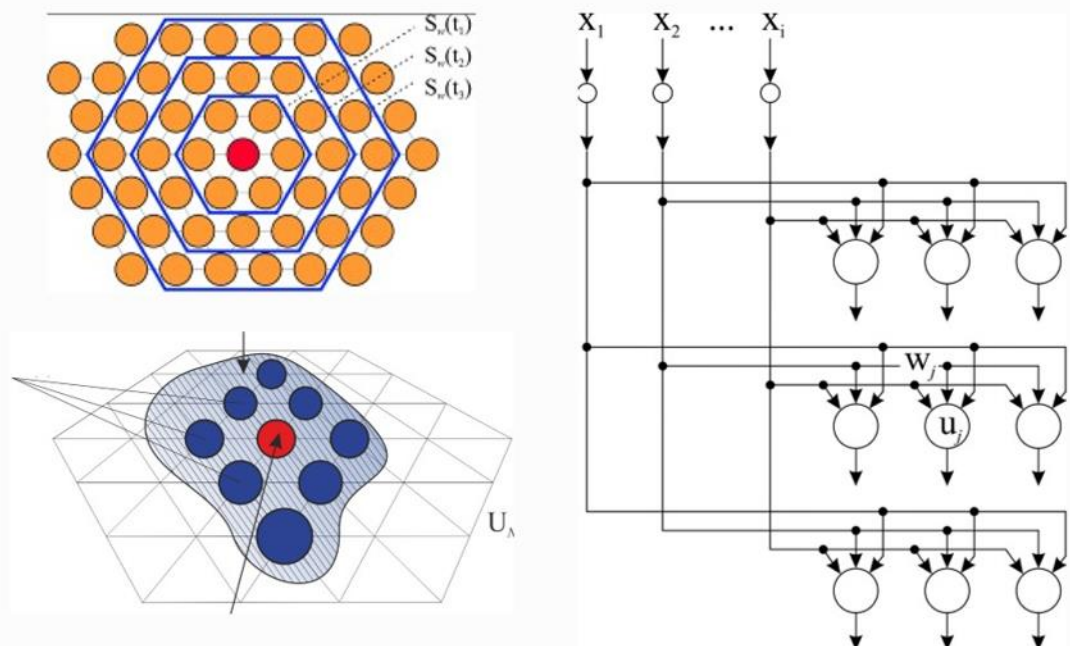
Класична модель карти Кохонена

3

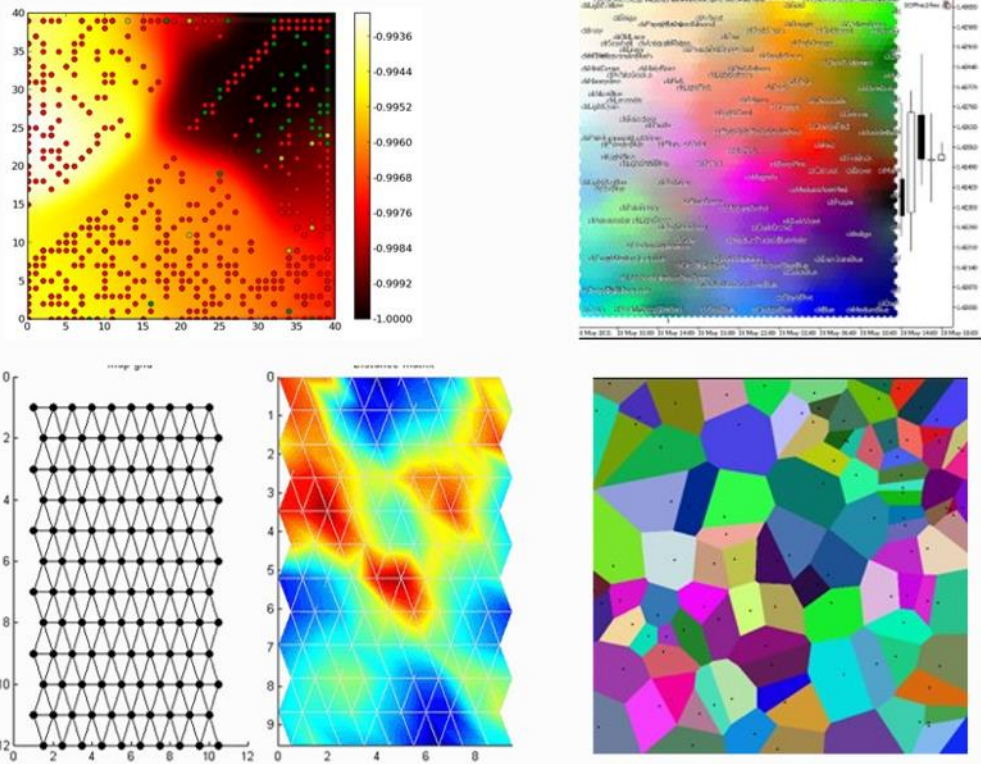


Класичний алгоритм навчання карти Кохонена

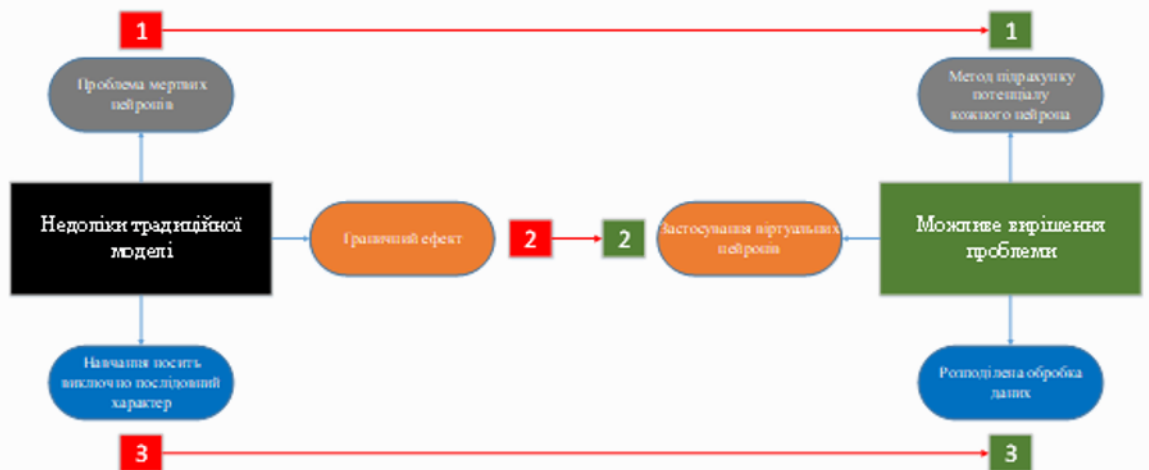
4



Побудова візуальних топографічних карт

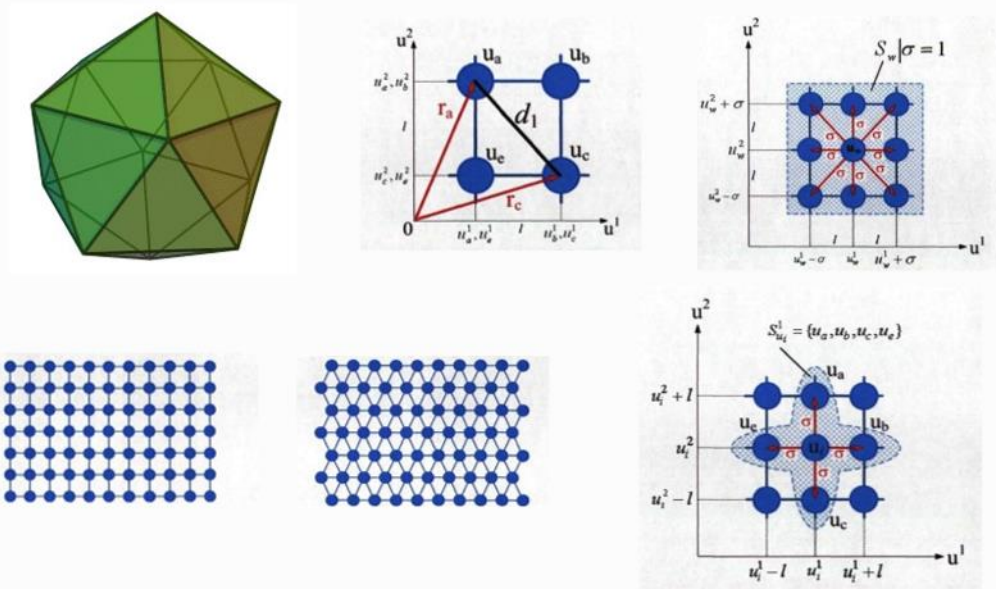


Проблеми застосування класичної моделі карти Кохонена



Метод модифікації класичного апарату карт Кохонена 7

Кохонена



Моделі карт Кохонена з замкнутою решіткою 8

$$u_{\max}^2 = \arg \max_{i \in [1, N]} \{u_i^2\}$$

$$u_{\min}^2 = \arg \min_{i \in [1, N]} \{u_i^2\}$$

$$\hat{S}_{u_i}^1 = S'_{u_i} \cup S''_{u_i} \cup S_{u_i}^1 | u_i \in U_N, i = 1, \dots, N$$

$$S'_{u_i} = \{u_j = (u_k^1, u_{\max}^2) \in U_N | u_k^1 \in [u_{\min}^1, u_{\max}^1], u_i^2 = u_{\max}^2, i \in [1, N]\}$$

$$S''_{u_i} = \{u_j = (u_k^1, u_{\min}^2) \in U_N | u_k^1 \in [u_{\min}^1, u_{\max}^1], u_i^2 = u_{\min}^2, i \in [1, N]\}$$

$$\tilde{S}_{u_i}^1 = S'_{u_i} \cup S''_{u_i} \cup S_{u_i}^1 | u_i \in U_N, i = 1, \dots, N$$

$$S'_{u_i} = \{u_j = (u_{\max}^1 - u_i^1, u_{\max}^2) \in U_N | u_i^2 = u_{\max}^2, i = 1, \dots, N\}$$

$$S''_{u_i} = \{u_j = (u_{\max}^1 - u_i^1, u_{\min}^2) \in U_N | u_i^2 = u_{\min}^2, i = 1, \dots, N\}$$

Ефективність навчання. Оцінка точності квантування мережі

9

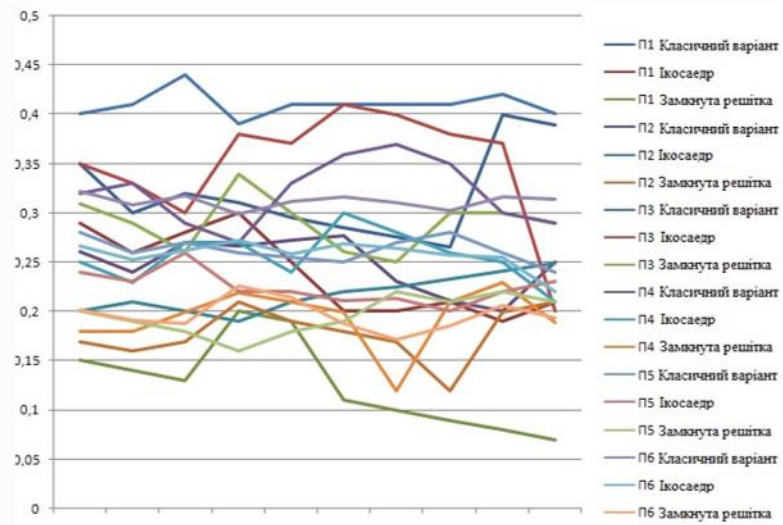
Кількість кроків навчання	Помилка квантування		
	Класична мережа, E_{q1}	Замкнута мережа, E_{q2}	Співвідношення E_{q1} / E_{q2}
10	0.094	0.041	2.293
30	0.159	0.037	4.251
50	0.155	0.036	4.270
100	0.136	0.036	3.747
200	0.119	0.036	3.324

Ефективність навчання. Ентропія мережі як оцінка якості навчання

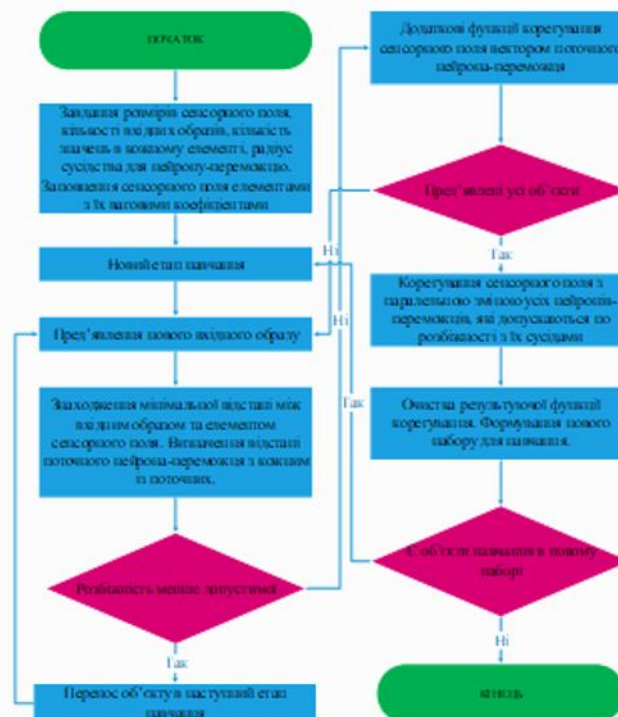
10

Кількість кроків навчання	Ентропія		
	Класична мережа, H_1	Замкнута мережа, H_2	Відхилення, % $(H_2 / H_1 - 1) * 100\%$
10	6.811	9.516	2.293
30	6.311	9.588	4.251
50	7.247	9.617	4.270
100	8.049	9.632	3.747
200	8.632	9.606	3.324

Графік залежності від величини помилки відобраного 11 варіанта навчання для всіх параметрів

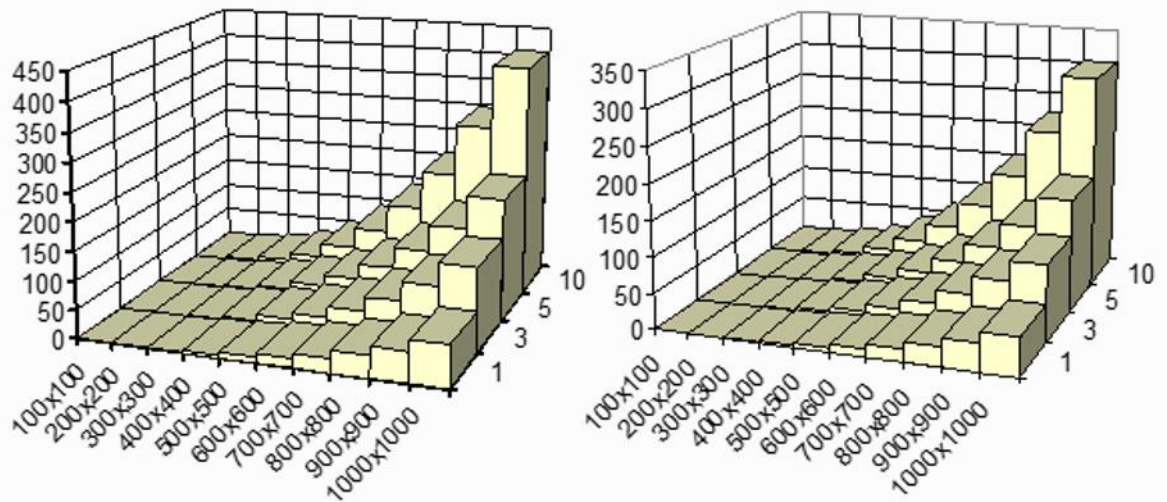


Блок-схема алгоритму навчання модифікованої карти 12 Кохонена



Порівняльний аналіз навчання традиційного та модифікованого варіантів навчання карт Кохонена

13



Висновки

14

Досліджено ефективність кластеризації даних за допомогою методів машинного навчання. Проведено дослідження методів кластеризації даних та методів машинного навчання. Особливу увагу приділено штучним нейронним мережам, зокрема картам Кохонена. Розроблено метод вдосконалення кластеризації даних за рахунок змін в решітці класичного апарату мереж Кохонена шляхом усунення недоліків мереж цього типу: граничного ефекту і появи "мертвих" нейронів. Проведено моделювання з використанням розробленого методу.