

# **ОСНОВИ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ, НЕЙРОННИХ МЕРЕЖ та ГЛИБОКОГО НАВЧАННЯ**

## **Модуль 3. Навчання без вчителя**

### **Лекція 3.5.**

#### **Кластеризація. Карти Кохонена (SOM)**

# Класичний AI / Класичний ML



Навчання без вчителя: Маємо великий набір даних. В цих даних є приховані закономірності.

**Задача** – знайти закономірності, наприклад, розбивши дані на певні групи чи кластери.

# Кластеризація

## Формально:

Маємо множину (вибірку)  $\mathbb{O}$  об'єктів  $o^{(j)}$ ,  $j = 1, 2, \dots, M$

Кожен об'єкт  $o^{(j)}$  має сукупність характеристик - ознак  $x_i^{(j)}$ ,  $i = 1, 2, \dots, N$  з множини  $\mathbb{X}$ .

Передбачається, що є множина  $\mathbb{C}$  класів (кластерів)  $c^{(k)}$ ,  $k = 1, 2, \dots, K < M$  (іноді  $K$  відомо, іноді – невідомо).

Але (на відміну від класифікації)!

**належність об'єкту  $o^{(j)}$  до класу  $c^{(k)}$  - невідома.**

# Кластеризація

Визначена деяка метрика  $d(o^{(j)}, o^{(i)})$  – відстань від між об'єктом  $o^{(j)}$  та об'єктом  $o^{(i)}$ .

**Завдання:** розбити вибірку  $o^{(j)}$ ,  $j = 1, 2, \dots, M$  на непересічні підмножини – кластери так, щоб кожен кластер складався з об'єктів, близьких по метриці  $d(., .)$ , а об'єкти різних кластерів істотно відрізнялися. При цьому кожному об'єкту  $o^{(j)}$  приписується відповідний кластер – клас  $c^{(k)}$ .

# Самоорганізаційна карта Кохонена

☹ Далі буде ☹

**The END**

**Модуль 3. Лекція 05.**