

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”
Інститут прикладної математики і фундаментальних наук

Кафедра прикладної математики

ЗВІТ

про виконання лабораторної роботи №3
із дисципліни
“Математичні основи штучного інтелекту”

Виконав: студент групи
ПМ-32, Шеремета Данило
Прийняв: доц. каф.
Пабірівський В. В.

Львів — 2023

Лабораторна робота №3

Тема: Кластеризація даних.

Мета: Навчитись реалізовувати програму алгоритму кластеризації та перевірити роботу алгоритму на тестовій множині даних.

Постановка задачі

Розробити програмну реалізацію алгоритму кластеризації та перевірити роботу алгоритму на тестовій множині даних.

Етапи виконання завдання

1. Вивчити із використанням запропонованих літературних джерел зміст задачі кластеризації та методи кластеризації даних.
2. Згенерувати тестову послідовність з N значень (для визначеності, можна покласти $N \geq 1000$), що є парами дійсних чисел на одиничному квадраті.
3. Реалізувати допоміжну функцію для обчислення міри віддалі.
4. Реалізувати допоміжну функцію для виконання алгоритму кластеризації за методом К-середніх (англ. K-means).
5. Реалізувати допоміжну функцію для виконання алгоритму кластеризації за будь-яким іншим (за вибором) методом (ієрархічний, С-середніх, мінімального покриваючого дерева, пошарової кластеризації, ART1, тощо).
6. Безпосередньо реалізувати кластеризацію даних двома методами та порівняти результати кластеризації.
7. Порівняйте кількість кластерів та якість кластеризації (можна просто оцінити середньозважене розмірів утворених кластерів згідно заданої міри віддалі для кожного із методів).

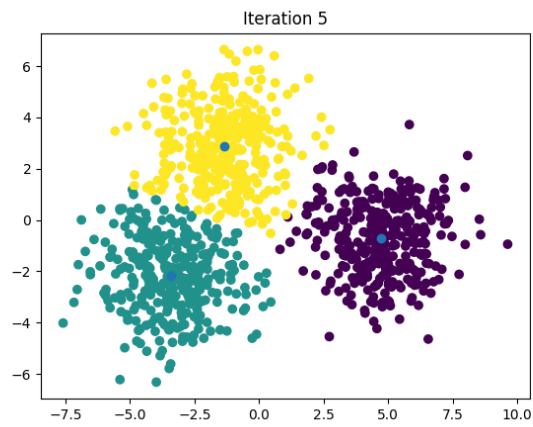
Посилання на GitHub:

https://github.com/DanSheremeta/mbai_labs/tree/main/Lab3

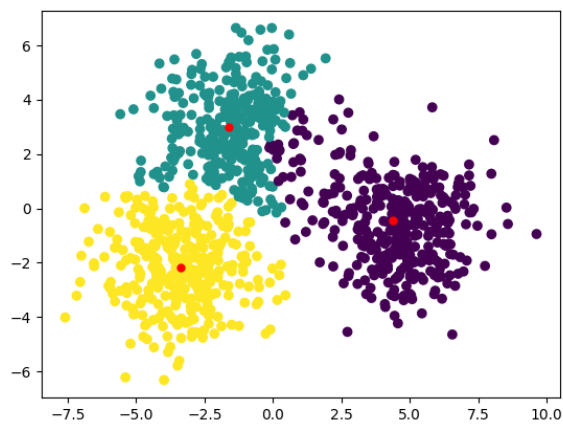
Порівняння двох методів

Кількість кластерів = 3, кількість точок = 1000

Метод К-середніх

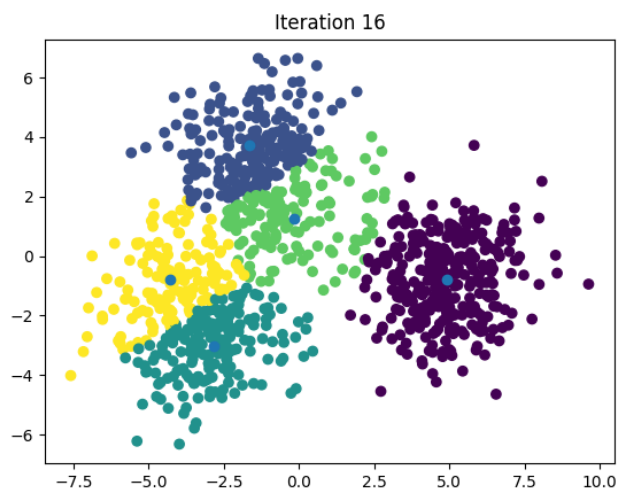


Метод ієрархії

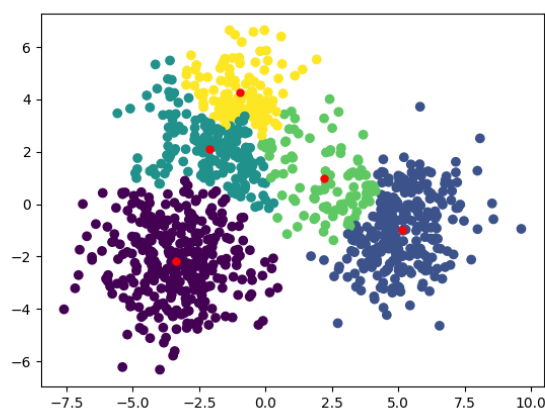


Кількість кластерів = 5, кількість точок = 1000

Метод К-середніх



Метод ієрархії



Висновок

Якщо порівнювати два методи для однакових даних можна дійти висновку, що їхні результати схожі. Метод ієрархії працює швидше оскільки йому не потрібно як в методі К-середніх використовувати ітерації.