

**Міністерство освіти і науки України
Національний технічний університет України
“Київський політехнічний інститут імені Ігоря Сікорського”
Фізико-технічний інститут**

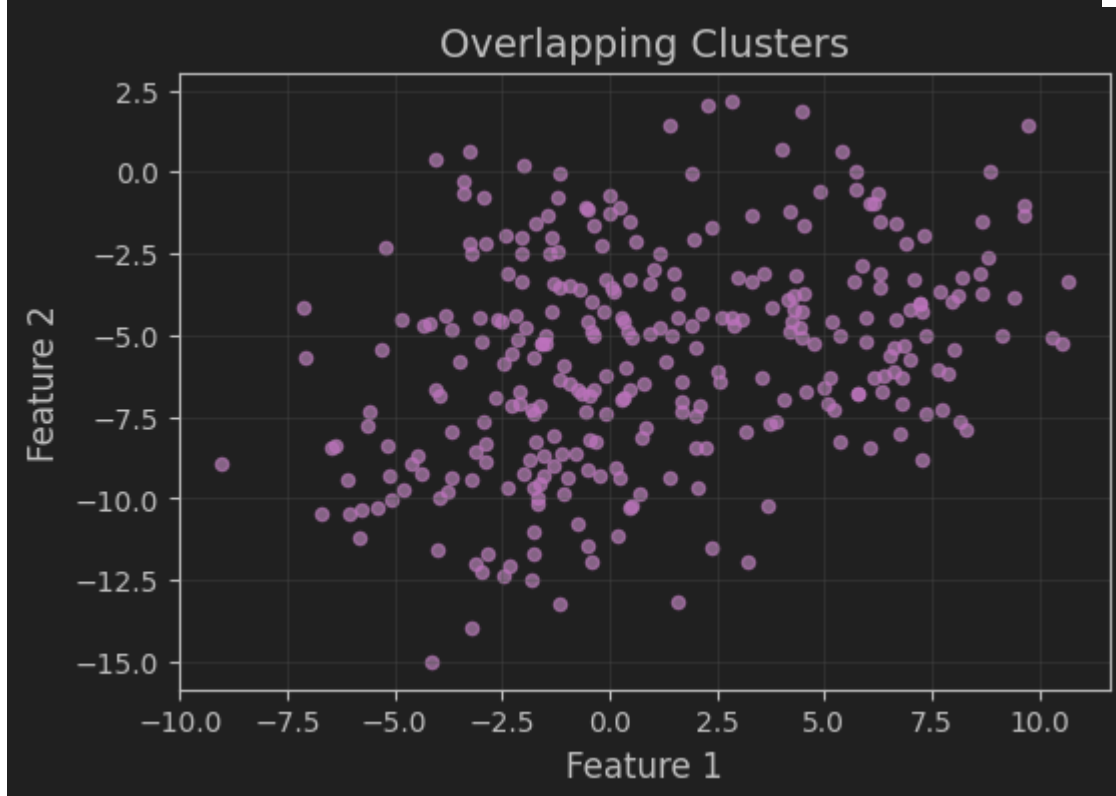
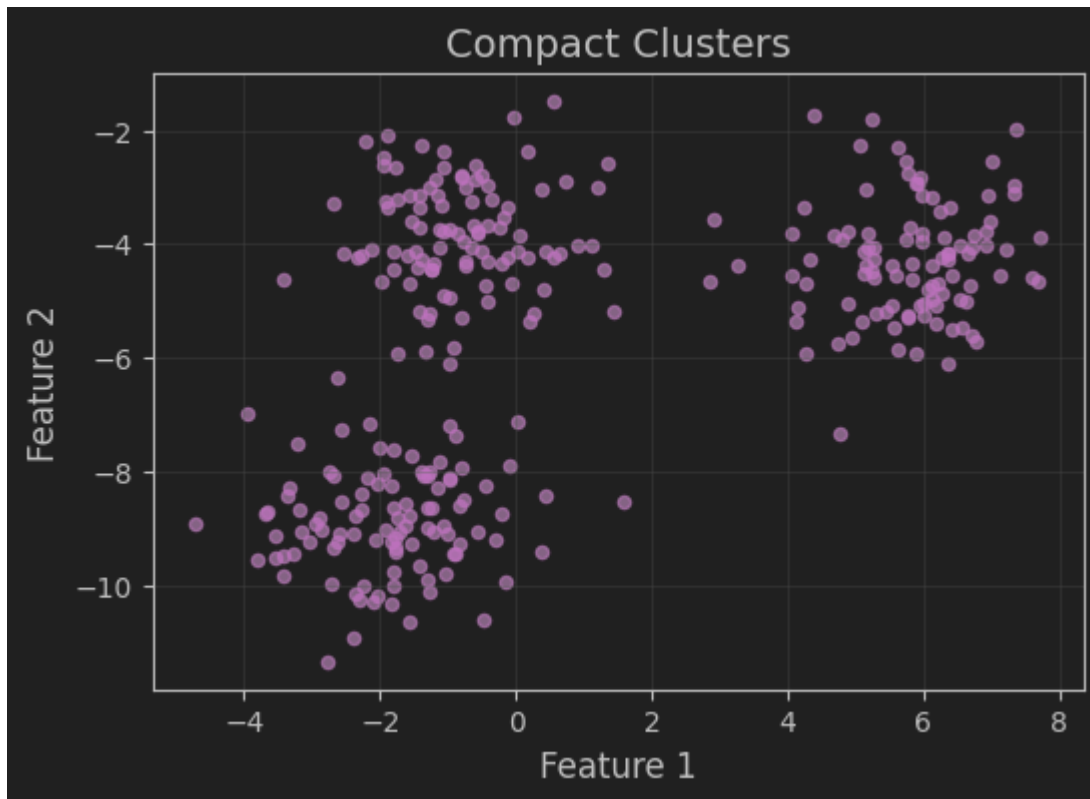
Методи обробки та розпізнавання даних
Лабораторна робота №3

Виконав:
студент ФІ-31мн
Шевченко Ю.
Перевірив:
Юзефович В. В.

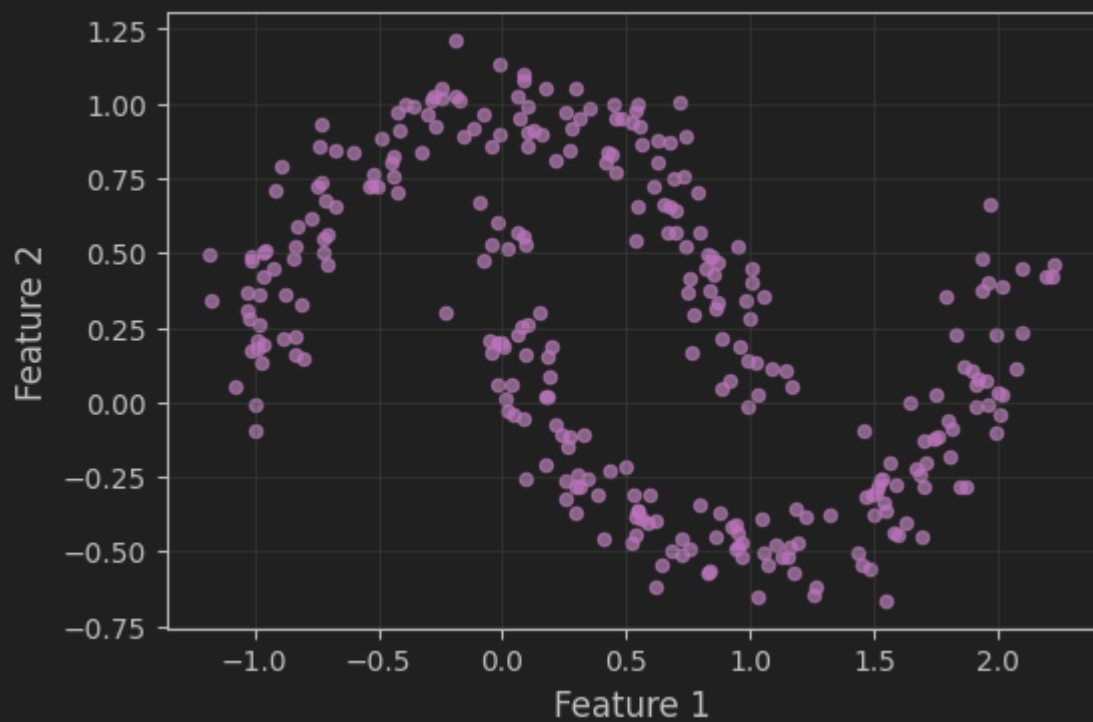
Лабораторна робота №3. Застосування методів кластеризації для обробки даних та зображень

Хід роботи:

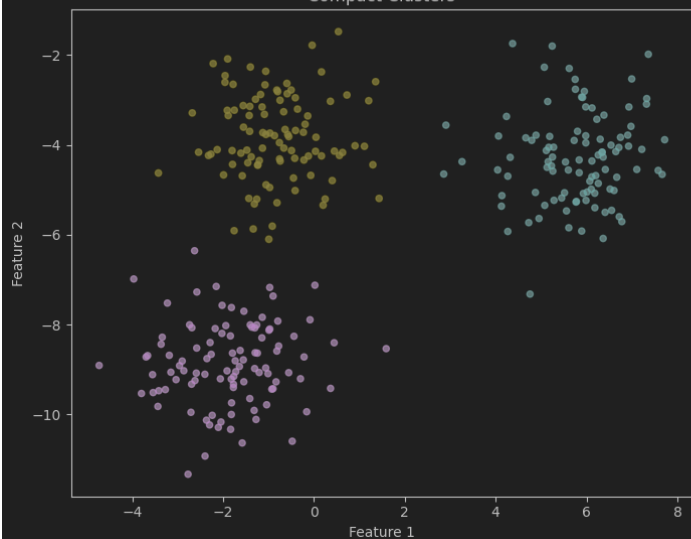
1.



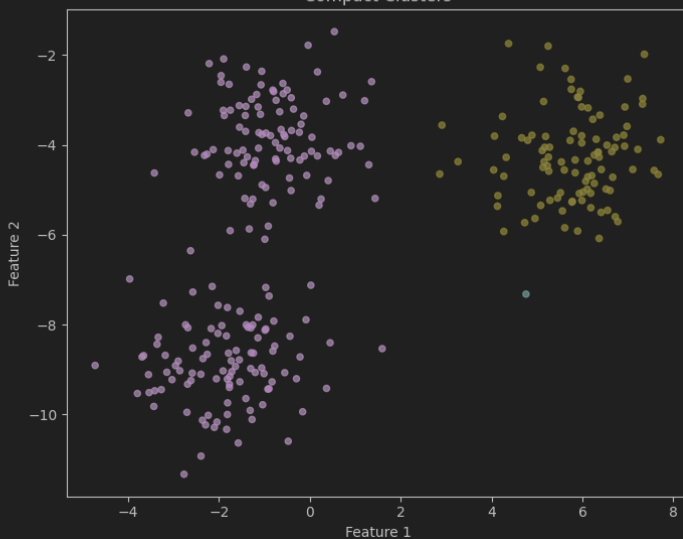
Non-linear Clusters



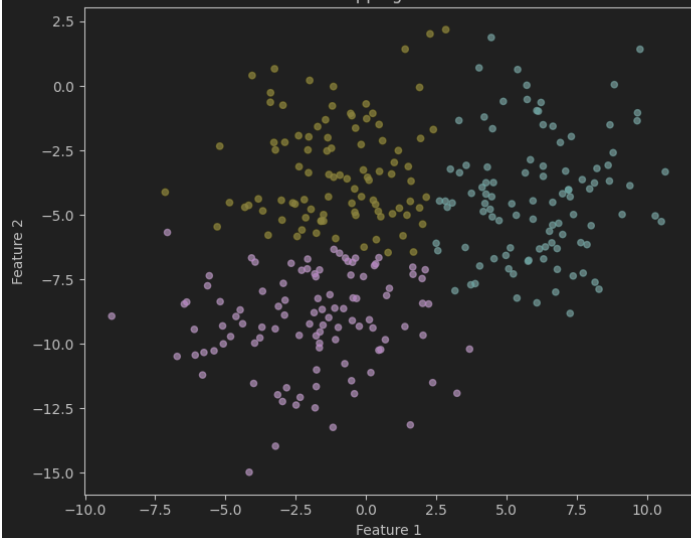
K-Means Clustering
Compact Clusters



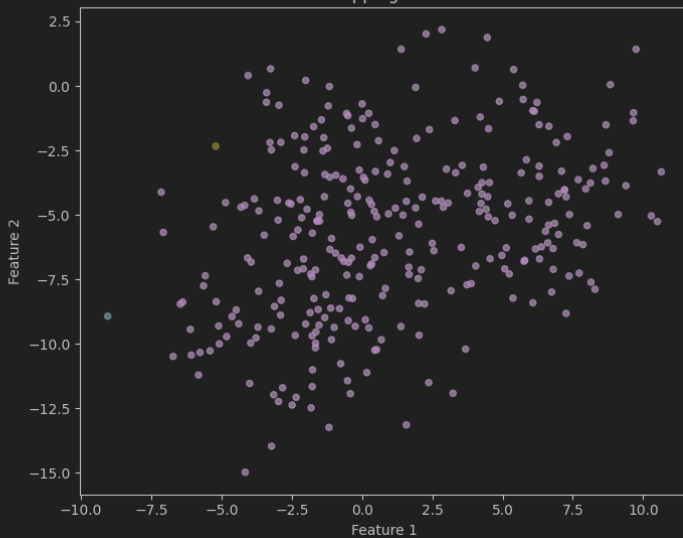
Agglomerative Clustering
Compact Clusters



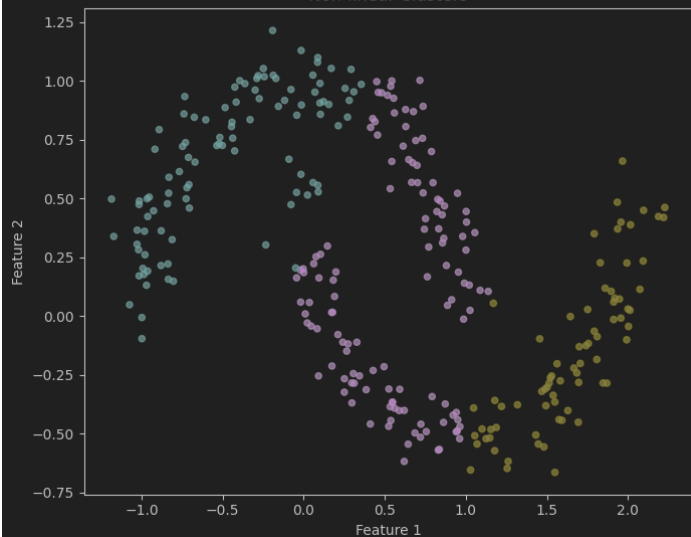
K-Means Clustering
Overlapping Clusters



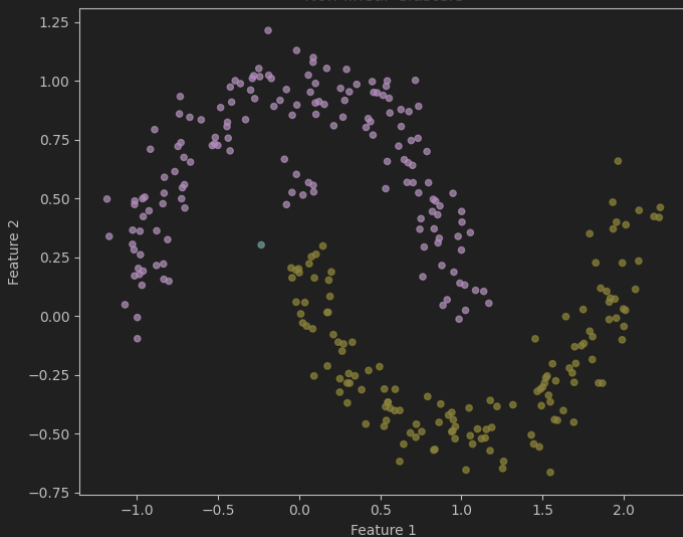
Agglomerative Clustering
Overlapping Clusters



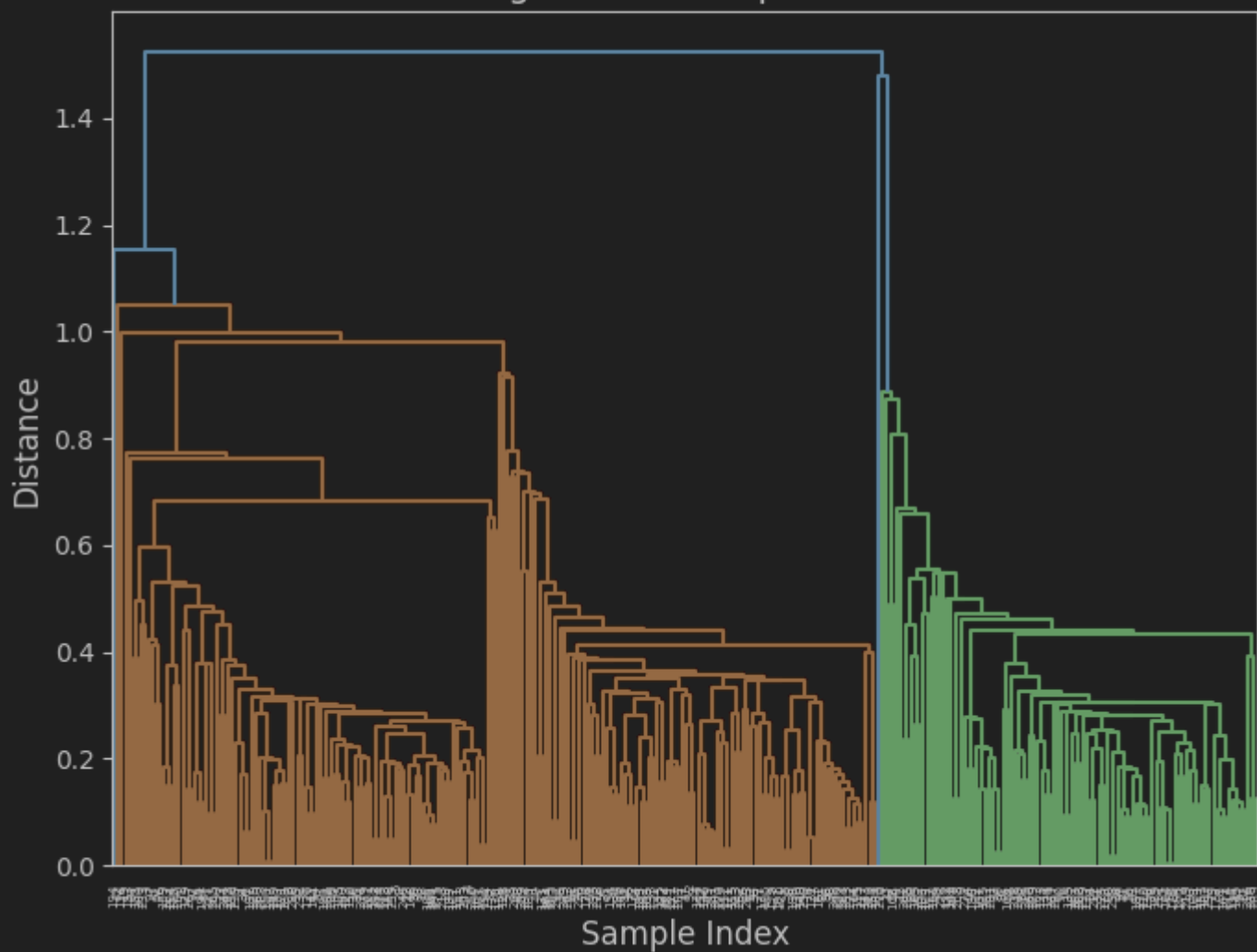
K-Means Clustering
Non-linear Clusters



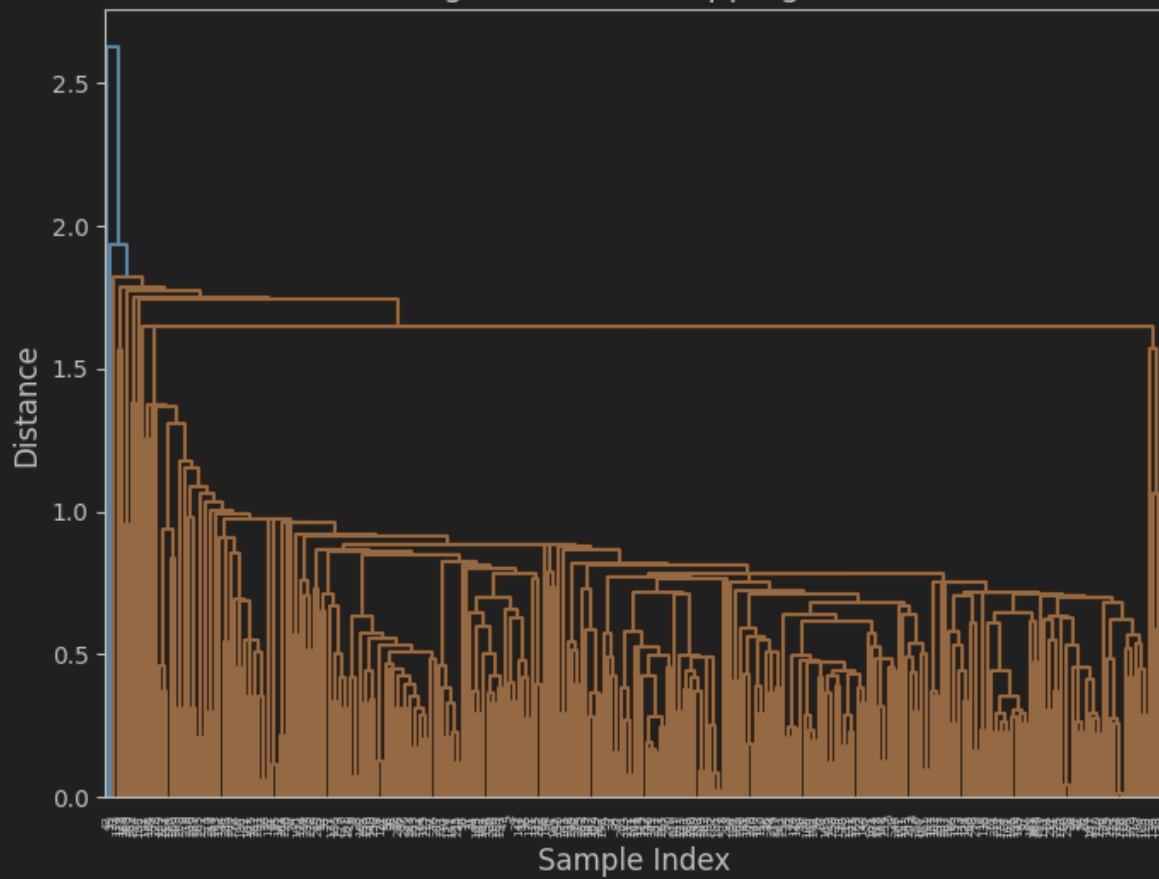
Agglomerative Clustering
Non-linear Clusters



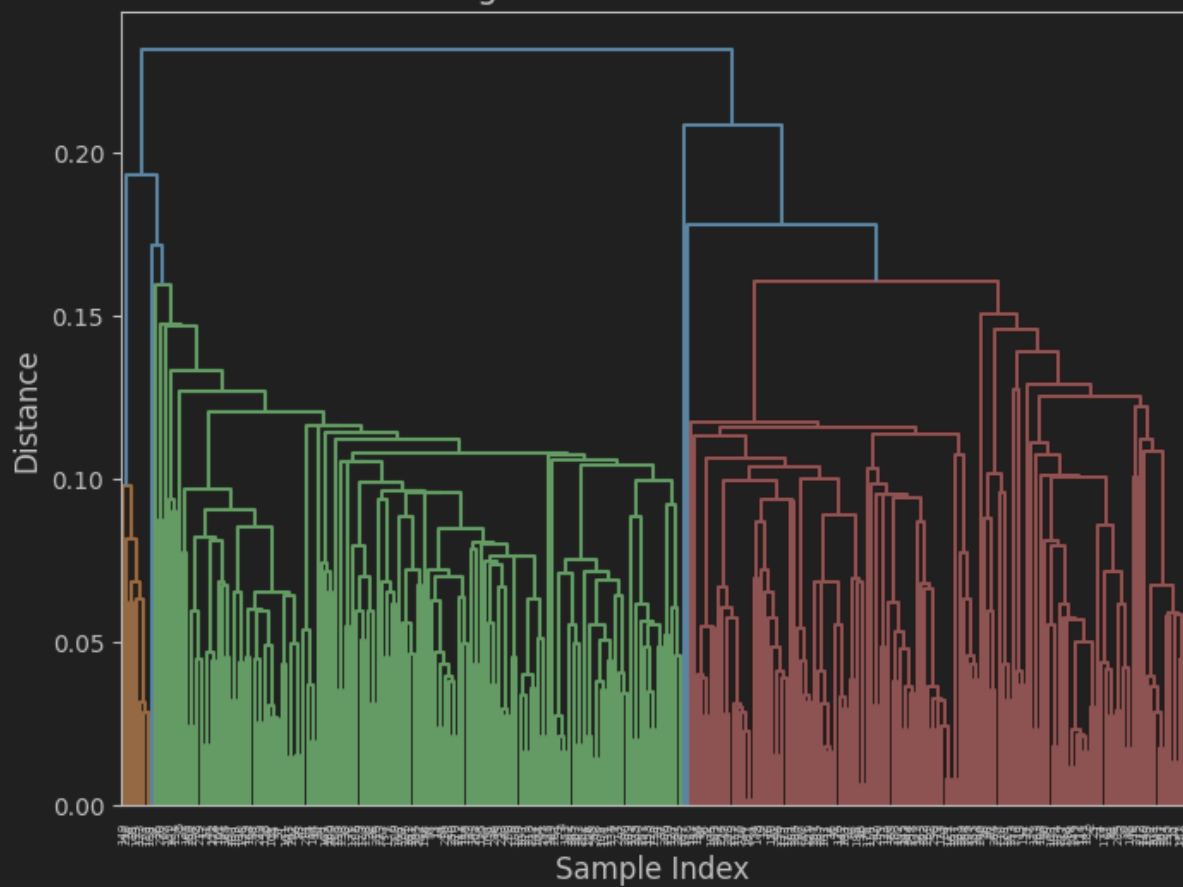
Dendrogram for Compact Clusters



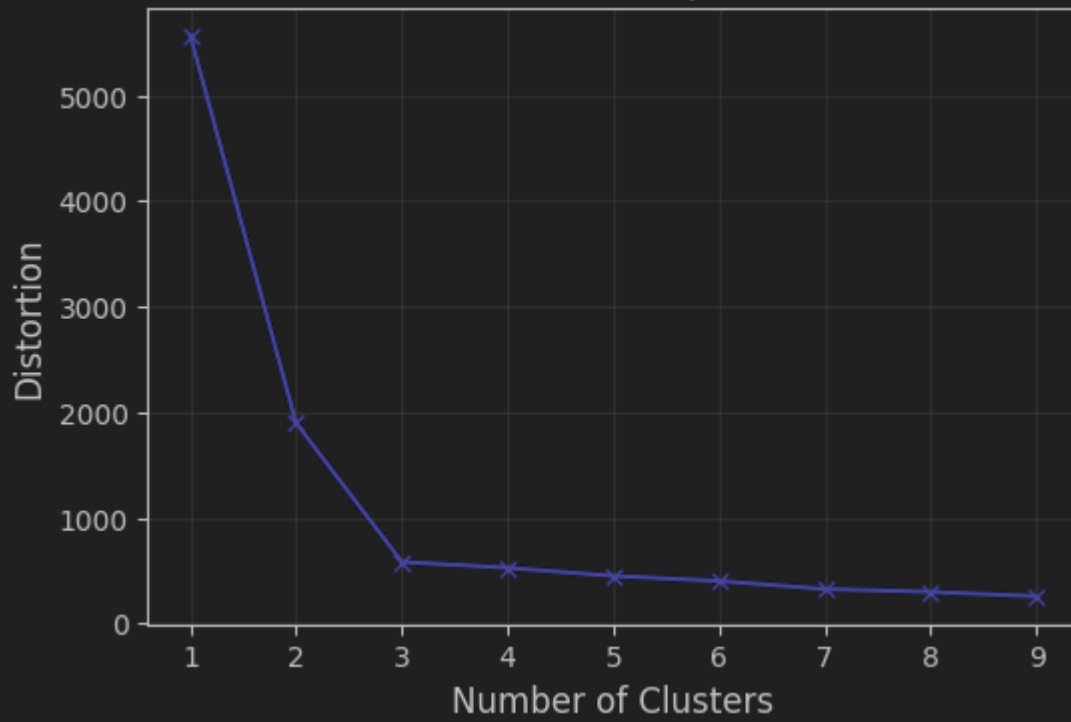
Dendrogram for Overlapping Clusters



Dendrogram for Non-linear Clusters

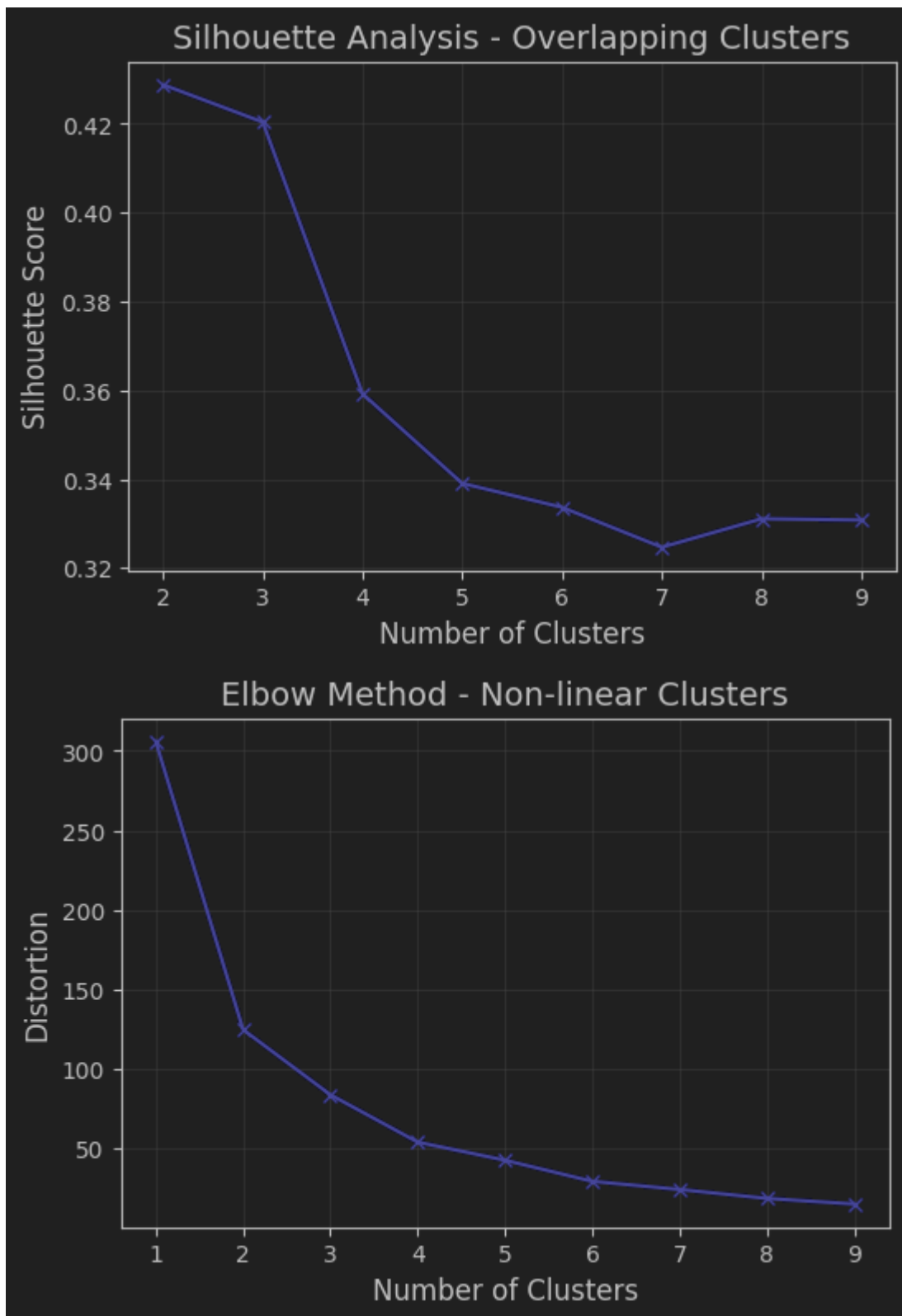


Elbow Method - Compact Clusters



Silhouette Analysis - Compact Clusters





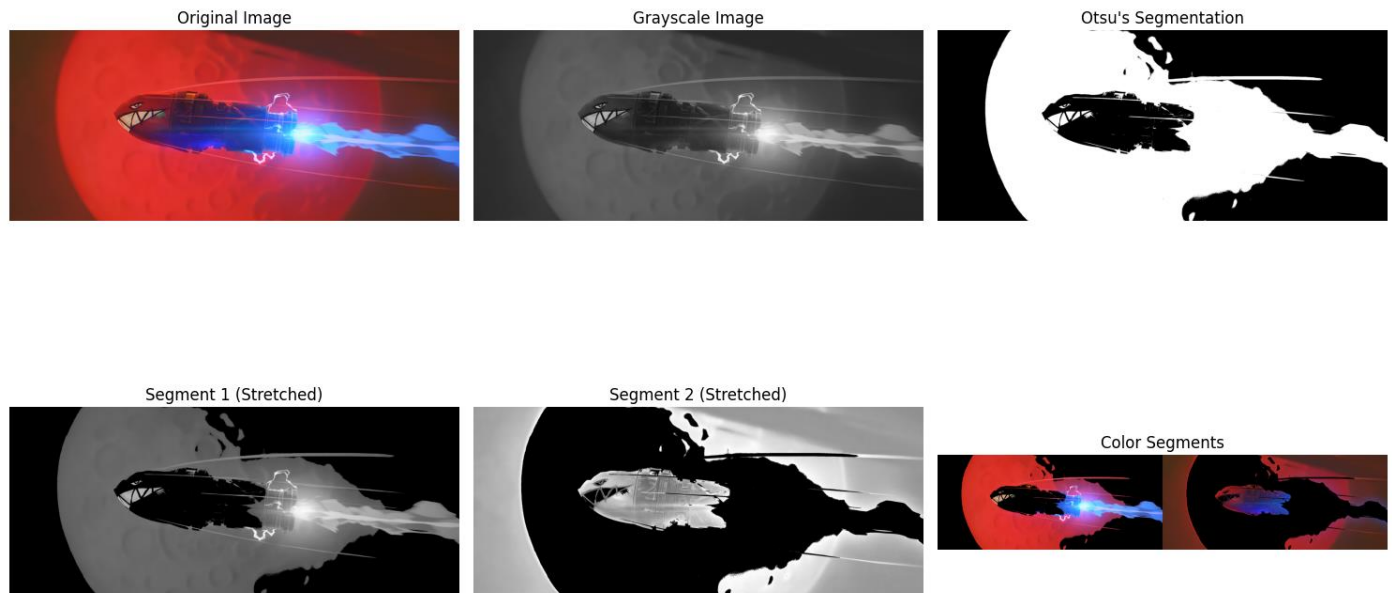
Результати показують, що метод KMeans є ефективнішим для компактних кластерів завдяки його здатності добре ідентифікувати чітко розділені групи об'єктів. Його силуетний коефіцієнт для таких кластерів найвищий. Водночас для випадків з перекриванням або нелінійним розподілом кластерів KMeans працює гірше, оскільки його основне припущення — сферичність кластерів — не відповідає цим розподілам.

Метод агломеративної кластеризації, орієнтований на пошук зв'язності між об'єктами, показав нижчі значення силуетного коефіцієнта в усіх сценаріях. Особливо він погано

справляється з перекриваннями та нелінійними розподілами, що свідчить про складність адаптації цього методу до більш складних структур.

Таким чином, KMeans є кращим вибором для структурованих, компактних даних, тоді як агломеративна кластеризація обмежена в точності для складніших типів розподілу.

2.



Метод Оцу продемонстрував свою ефективність для сегментації зображень на основі яскравості. Оптимальний поріг, визначений цим методом, дозволив чітко розділити об'єкти на два сегменти, що підтверджується якістю отриманих бінарних зображень. Розтягування гістограм яскравості для окремих сегментів дозволило покращити їхню візуалізацію, а отримані кольорові сегменти показали добру відповідність об'єктам на вхідному зображенні. Це підтверджує, що метод Оцу є ефективним для зображень з чіткими яскравими контрастами.

Метод кластеризації k-середніх, застосований для сегментації на основі кольору, продемонстрував гнучкість у виділенні контрастних об'єктів. Оптимальне значення k дозволило досягти найкращого візуального результату, розділивши об'єкти за кольоровими характеристиками. Проте ефективність методу залежить від попереднього вибору кількості кластерів і характеру кольорового розподілу у зображенні.

Загалом, метод Оцу є більш придатним для простих задач сегментації на основі яскравості, особливо коли об'єкти мають високий контраст. Водночас k-середніх добре підходить для задач, де колір є ключовим фактором, але потребує додаткового налаштування параметрів для досягнення оптимальних результатів.

Висновки:

У цій роботі проведено кластеризацію даних та сегментацію зображень, що дозволило оцінити ефективність різних методів залежно від специфіки завдань. Методи кластеризації, такі як k-середніх та агломеративна кластеризація, продемонстрували свої переваги та обмеження залежно від структури даних. Для сегментації зображень метод Оцу виявився ефективним у задачах із чіткими яскравими контрастами, тоді як кластеризація k-середніх забезпечила якісне розділення об'єктів за кольором. Результати показали, що вибір методу має залежати від характеру даних та вимог до точності й деталізації сегментації чи кластеризації.