

20 - 2 인공지능 과제 2

# Hierarchical Clustering Problem

---

조교 조충현, 이봉석  
johnart@hanyang.ac.kr  
tjr4090@hanyang.ac.kr

# 과제 개요

- Hierarchical Agglomerative Clustering 기법을 이용하여 좌표평면상의 점들 clustering
- 세부사항
  - 총 3가지의 좌표평면의 점들에 대한 데이터를 각각 3개의 군집으로 묶일 수 있게 하는 similarity level을 구하는 것이 목표
  - 좌표평면상의 점은 각각  $n$ 개가 주어진다.
  - 점들 사이의 similarity는 Cosine Similarity로 정의한다.
  - 좌표평면상의 점들을 single-link, complete-link, group average-link clusterings 방법을 통해 각 방법마다 3개의 군집으로 나눠지게 하는 similarity level 범위를 구한다.

## 결과물 예시

- 입력 : k번째 좌표평면(k) / 점의 개수(n) / 점들의 좌표 (x, y)  
(n <= 500, -1000 <= x, y <= 1000)

1 5

5,3

10,15

15,12

24,10

83,23

# 결과물 예시

- 출력 : k번째 좌표평면(k) / 각 방법의 clusters, similarity level 범위

1

---

single

clusters: [(10, 15)], [(24, 10), (83, 23)], [(5, 3), (15, 12)]

span: 0.99099243, 0.98941491

---

complete

clusters: [(10, 15)], [(24, 10), (83, 23)], [(5, 3), (15, 12)]

span: 0.99099243, 0.91933251

---

average

clusters: [(10, 15)], [(24, 10), (83, 23)], [(5, 3), (15, 12)]

span: 0.99099243, 0.95839063

# 과제 조건

- 환경

- 프로그래밍 언어 : **Python 3.x**
- OS : **Windows 10**
- 외부 라이브러리 사용 **불가**(내장함수만 사용 가능)
- 보고서 : **MS Word**

- 제출 사항

- 파이썬 파일 : 본인학번\_assignment\_2.py (파이썬 **파일 한 개**) (70%)
- 결과 보고서 : 본인학번\_assignment\_2.docx
  - 코드 설명 (10%)
  - 함수 설명 (10%)
  - 실험 결과 설명 (10%)
- 출력파일 (3개) : **CoordinatePlane\_N\_output.txt** (N은 N번째 Coordinate Plane)

# 주의 사항

- 파일명 반드시 준수.
- 파일은 **Blackboard**에 올려주세요.
- 제출 기한 : **2020.11.20 (23:59)**
- 추가 제출 기한 **없음**.
- 점수 비중 : **코드 70% 보고서 30%**

# 주의 사항

- 아래 예시와 같은 입출력 형식을 반드시 준수.
- 코드와 같은 경로에 input.txt와 output.txt 관리.

**입력** input.txt 에서 read

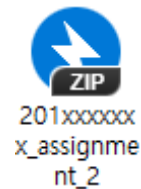
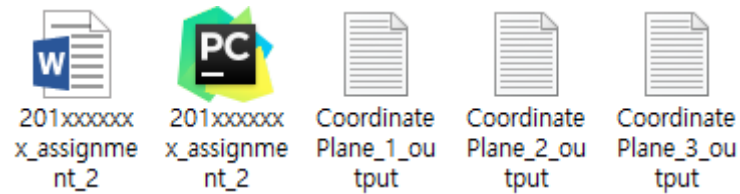
```
1 5
5,3
10,15
15,12
24,10
83,23
```

**출력** output.txt 로 write

```
1
---
single
clusters: [(10, 15)], [(24, 10), (83, 23)], [(5, 3), (15, 12)]
span: 0.99099243, 0.98941491
---
complete
clusters: [(10, 15)], [(24, 10), (83, 23)], [(5, 3), (15, 12)]
span: 0.99099243, 0.91933251
---
average
clusters: [(10, 15)], [(24, 10), (83, 23)], [(5, 3), (15, 12)]
span: 0.99099243, 0.95839063
```

# 주의 사항

- 제출하는 파일들을 압축하여 BlackBoard에 올릴 것!





Thank you!

---