

Minggu ke-13

#### **Praktikum Clustering**

Entin Martiana, Ali Ridho Barakbah

Knowledge Engineering Laboratory

Department of Information and Computer Engineering

Politeknik Elektronika Negeri Surabaya

## Clustering dengan k-Means

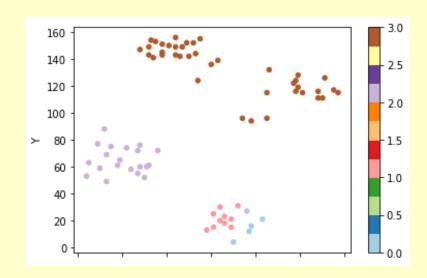
```
import pandas as pd
from sklearn.cluster import KMeans

dataset = pd.read_csv('ruspini.csv')
data = dataset.loc[:,['X', 'Y']]

clustering = KMeans(n_clusters=4, init="random", n_init=1)
clusters=clustering.fit_predict(data)

print('\nHasil clustering:\n', clusters)

data.plot(x='X', y='Y', kind='scatter', c=clusters, colormap='Paired')
```



### Clustering dengan Hierarchical Clustering

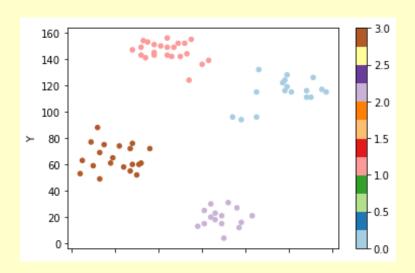
import pandas as pd from sklearn.cluster import AgglomerativeClustering

dataset = pd.read\_csv('ruspini.csv')
data = dataset.loc[:,['X', 'Y']]

clustering=AgglomerativeClustering(n\_clusters=4, linkage='average') clusters=clustering.fit\_predict(data)

print('\nHasil clustering:\n', clusters)

data.plot(x='X', y='Y', kind='scatter', c=clusters, colormap='Paired')





# Cluster Analysis dengan SSE

```
import pandas as pd
from sklearn.cluster import KMeans

dataset = pd.read_csv('ruspini.csv')
 data = dataset.loc[:,['X', 'Y']]

clustering = KMeans(n_clusters=4, init="random", n_init=1)
 clusters=clustering.fit_predict(data)

print('\nHasil clustering:\n', clusters)

print('\nSSE = :\n', clustering.inertia_)
```

# Assignment #

- 1. dataset ← heart.csv, dan tampilkan
- 2. data ← normalisasi dengan min-max(0-1)
- 3. cluster ← lakukan clustering pada data menggunakan K-means (k=2)
- cluster ← lakukan clustering pada data dengan Single, Average,
   Complete Linkage, dengan k=2
- 5. Lakukan untuk setiap jumlah atribut di langkah ke-4 : cluster\_i[1-10], cluster\_val[1-10] ← lakukan clustering pada data dengan atribut yang paling berpengaruh dengan K-Means, dengan k=3, sebanyak 10 kali. Setiap kali selesai clustering, lakukan cluster analysis dengan SSE.
- 6. cluster ← ambil cluster\_i yang mempunyai cluster\_val terkecil

# Pengumpulan Tugas

- Buatlah coding dengan Bahasa pemrograman/tools apapun untuk semua assignment
- Buatlah laporan dalam slide ppt. Laporan terdiri dari screenshot coding dan hasil running untuk setiap assignment.
- Simpan laporan dalam file pdf dengan format penamaan: AID4ITB\_M13\_NRP\_namadepan.pdf
- Upload file tersebut ke ETHOL
- Deadline upload: Jum'at, 21 Mei 2023 pk. 23.00