K-MEANS CLUSTERING



**LAPORAN PRAKTIKUM  
Disusun untuk Memenuhi Tugas Laporan Praktikum Pertemuan 3  
Mata Kuliah Pembelajaran Mesin**

**DISUSUN OLEH:  
LINGGAR MARETVA CENDANI  
24060117120031**

**PROGRAM STUDI STRATA 1 INFORMATIKA  
DEPARTEMEN ILMU KOMPUTER/INFORMATIKA  
FAKULTAS SAINS DAN MATEMATIKA  
UNIVERSITAS DIPONEGORO  
SEMARANG  
2019**

# BAB I RUMUSAN MASALAH DAN TUJUAN

## 1.1 Rumusan Masalah

1. Bagaimana cara melakukan clustering menggunakan sample dataset?
2. Bagaimana cara melakukan clustering menggunakan generate dataset?
3. Bagaimana cara melakukan evaluasi hasil clustering?

## 1.2 Tujuan

1. Mengetahui cara melakukan evaluasi hasil clustering.
2. Mengetahui cara melakukan evaluasi hasil clustering.
3. Mengetahui cara melakukan evaluasi hasil clustering.

# BAB II DASAR TEORI

## 2.1 Scikit-learn

## Scikit-learn adalah pustaka Python open source yang mengimplementasikan serangkaian pembelajaran mesin, pra-pemrosesan, validasi silang, dan algoritma visualisasi menggunakan antarmuka terpadu.

## 2.2 Matplotlib

Matplotlib adalah library Python 2D yang dapat menghasilkan plot dengan kualitas tinggi dalam berbagai format dan dapat digunakan di banyak platform.

## Matplotlib dapat digunakan sebagai pembuat grafik dalam berbagai platform, seperti Python dan Jupyter. Grafik yang dapat dibuat beragam, seperti grafik garis, batang, lingkaran, histogram, dsb.

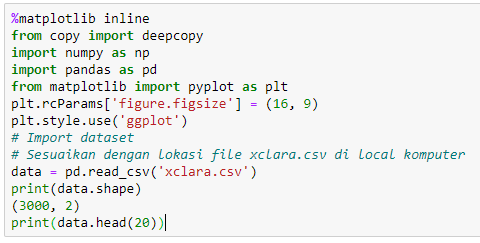
## 2.3 Numpy

NumPy adalah paket pemrosesan array tujuan umum. Ini menyediakan objek array multidimensi berkinerja tinggi, dan alat untuk bekerja dengan array ini. Ini adalah paket mendasar untuk komputasi ilmiah dengan Python. Ini berisi berbagai fitur termasuk yang penting ini seperti Objek array N-dimensi yang kuat, Fungsi-fungsi canggih (penyiaran), Alat untuk mengintegrasikan kode C / C ++ dan Fortran, Aljabar linier yang bermanfaat, transformasi Fourier, dan kemampuan angka acak.

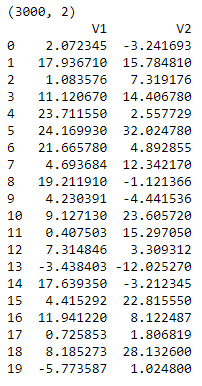
# BAB III PEMBAHASAN

## 3.1 Melakukan Impor Dataset

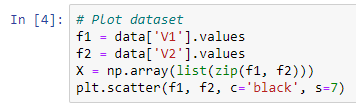
Dataset yang digunakan sebagai contoh adalah xclara.csv. Dataset tersebut terdiri atas 3000 record dan setiap record berisi 2 fitur. Secara natural dataset tersebut membentuk 3 cluster.



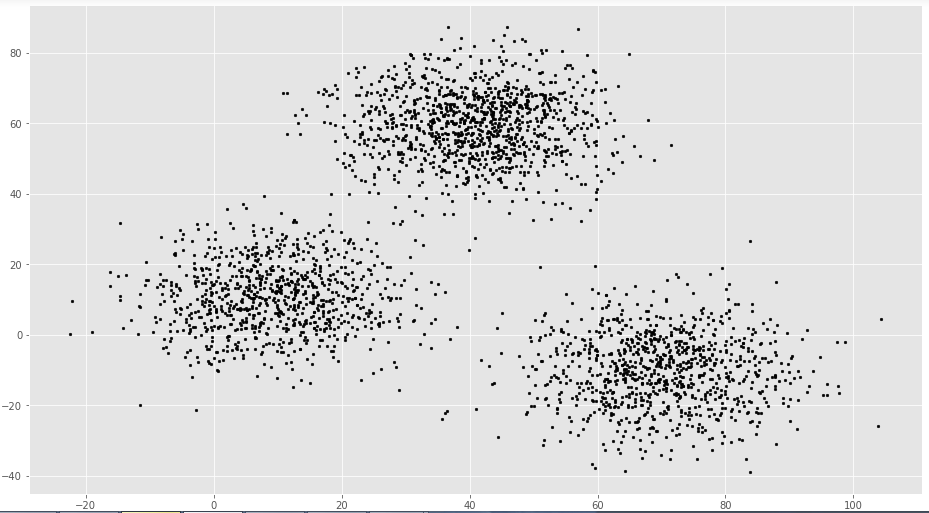
Maka akan ditampilkan data sebagai berikut.



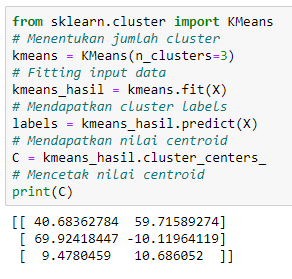
Kemudian dilakukan plot dataset.



Maka akan ditampilkan grafik sebagai berikut.



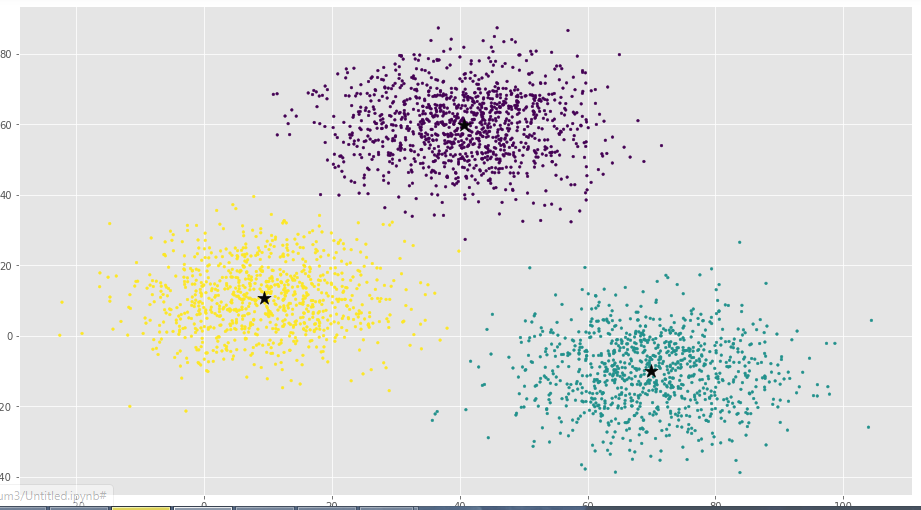
Kemudian dilakukan clustering



Kemudian dilakukan plot hasil Clustering.

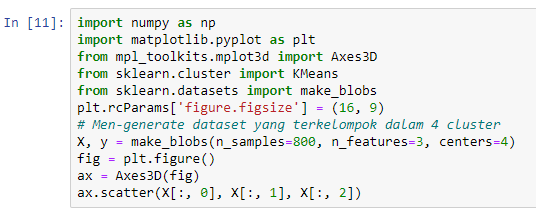


Dengan hasil sebagai berikut.

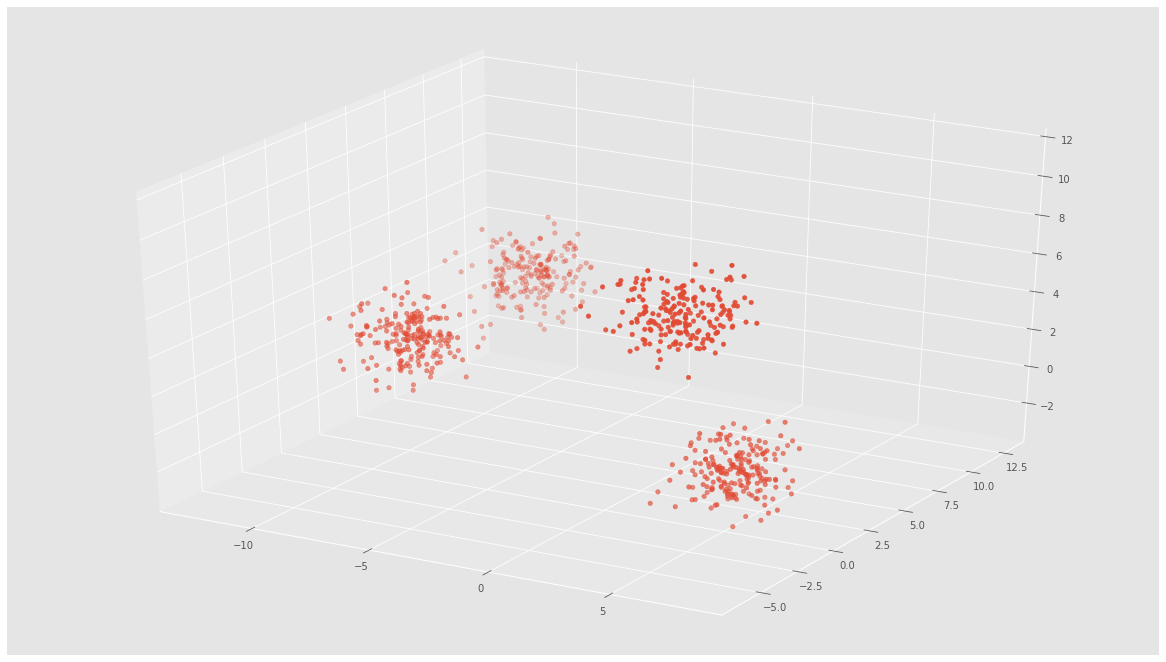


## 3.2 Melakukan Clustering Menggunakan Generate Dataset

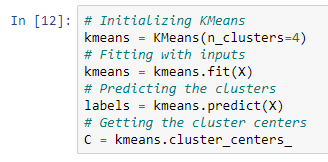
Dataset akan digenerate otomatis menggunakan make\_blobs. Pertama kita generate dataset.



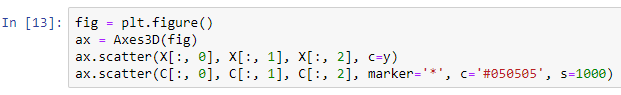
Maka hasilnya adalah sebagai berikut.



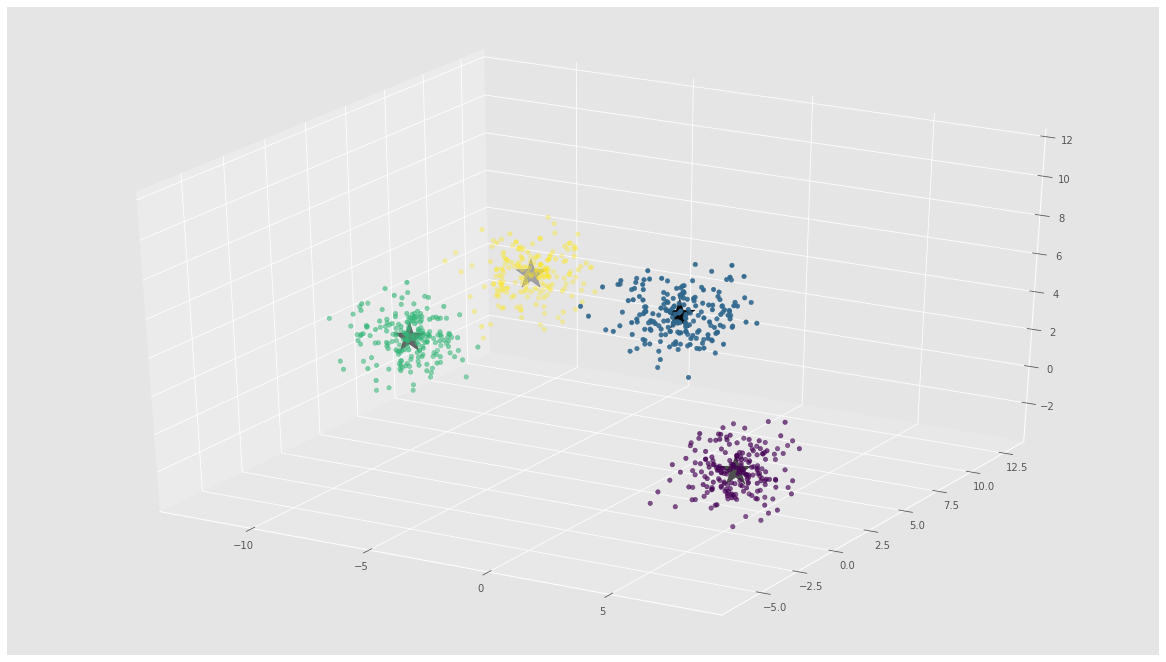
Lalu dilakukan clustering.



Lalu lakukan plot hasil clustering.

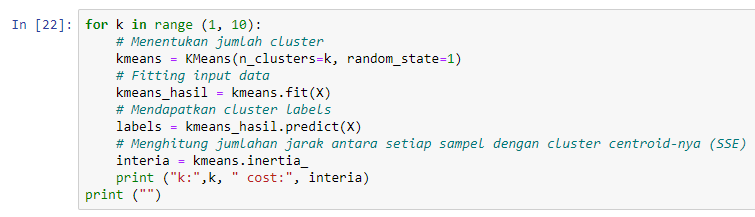


Maka hasilnya akan sebagai berikut.

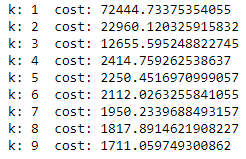


## 3.3 Melakukan Evaluasi Hasil Clustering

Selanjutnya, kita perlu melakukan evaluasi hasil clustering menggunakan beberapa kemungkinan nilai k dan menggunakan metric yang bersesuaian untuk clustering. Sebagai contoh kita akan mengevaluasi hasil clustering pada contoh yang pertama.



Maka hasilnya akan sebagai berikut



Kemudian lakukan evaluasi hasil cluster menggunakan silhouette coefficient.



Maka hasilnya akan sebagai berikut.



# BAB IV KESIMPULAN

## 1 Kesimpulan

Untuk melakukan K-Means clustering pada data yang kita punya kita perlu menyiapkan data sample dan juga data generate. Kemudian setelah dilakukan clustering. Kita juga perlu melakukan evaluasi pada clustering yang kita lakukan.

# 

# DAFTAR PUSTAKA

Informatics (28 Oktober 2019). "*Pertemuan 3 - K-Means Clustering*". Makalah disajikan dalam Praktikum Pembelajaran Mesin di laboratorium komputer gedung E Universitas Diponegoro. Semarang, 28 Oktober 2019.

GeeksforGeeks (8 April 2017). "*Learning Model Building in Scikit-learn : A Python Machine Learning Library*". Retrieved 4 November 2019. from GeeksforGeeks : https://www.geeksforgeeks.org/learning-model-building-scikit-learn-python-machine-learning-library/

GeeksforGeeks (10 Juli 2018). "*NumPy in Python | Set 1 (Introduction)*". Retrieved 4 November 2019. from GeeksforGeeks : https://www.geeksforgeeks.org/numpy-in-python-set-1-introduction/