



# Laporan Tugas 2

## *Machine Learning*

### ***K-Means***

Disusun oleh :

Odia Pratama      1301154405

IF-39-13

Universitas Telkom

Bandung

2018

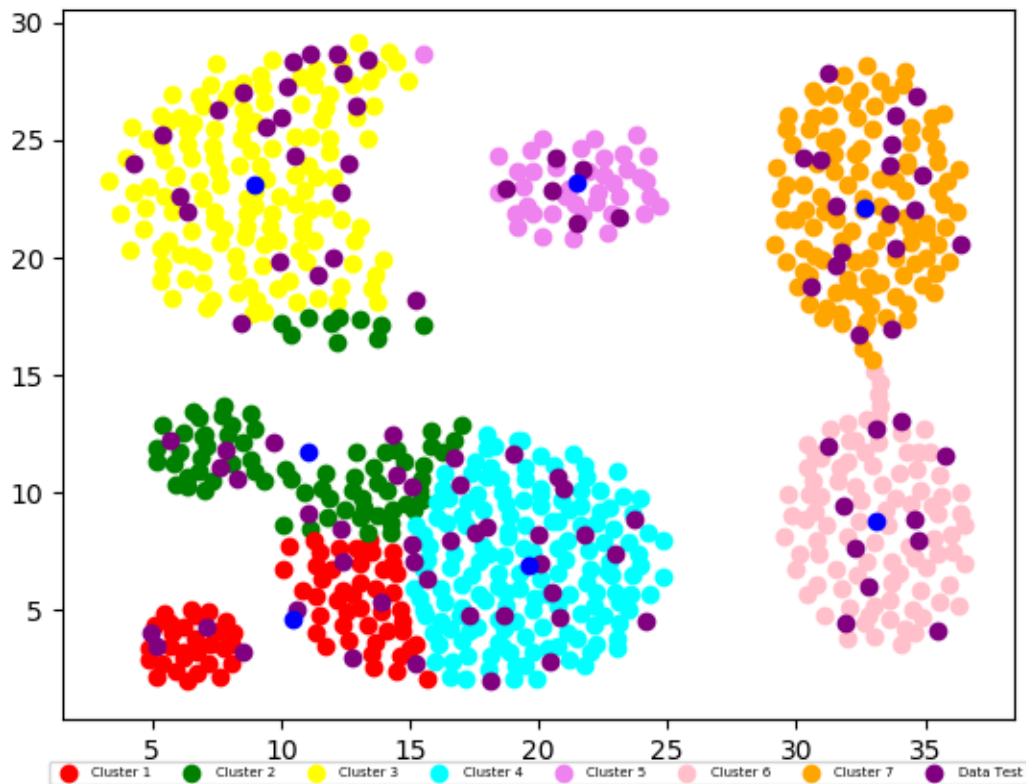
## Analisis Masalah

Pada tugas ini disediakan 688 *data training* dan 100 *data testing*. Pada data yang telah sediakan akan dilakukan proses klasifikasi pada data training terlebih dahulu untuk menentukan label klasifikasi pada *data training*. Pada data testing yang tersedia dilakukan proses untuk menentukan jarak titik centroid ke titik pada data testing untuk menetapkan *cluster* tiap titik yang ada, sehingga pada data testing akan diuji untuk setiap titik untuk menentukan *cluster* pada titik tersebut.

## Desain



## Evaluasi Hasil Eksperimen



Pada pemrograman yang dibangun berdasarkan desain metode k-means, hasil dari running program menghasilkan 7 cluster dengan warna (*red, green, yellow, cyan, violet, pink, orange*). Titik Centroid akhir yang merupakan penentu dari cluster tiap titik yang ada. Data testing pada gambar ditunjukkan dengan warna ungu, sehingga setiap titik dari data testing akan dihitung jaraknya ke setiap cluster, jarak titik yang paling dekat dengan centroid itulah yang menjadi hasil klasifikasi data.

Kesimpulan dari percobaan tugas ini adalah, untuk menentukan jumlah K dapat dilihat dari penyebaran data, dari data tersebut dapat dilihat pusat-pusat dari titik-titik yang ada. Atau dengan menggunakan perhitungan SSE (*Sum Square Error*), sehingga dapat dilihat dari grafik pada percobaan berbagai K secara berurut yang tidak mengalami penurunan secara signifikan.