

LAPORAN TUGAS 2.1
PEMBELAJARAN MESIN
k-Mean Clustering

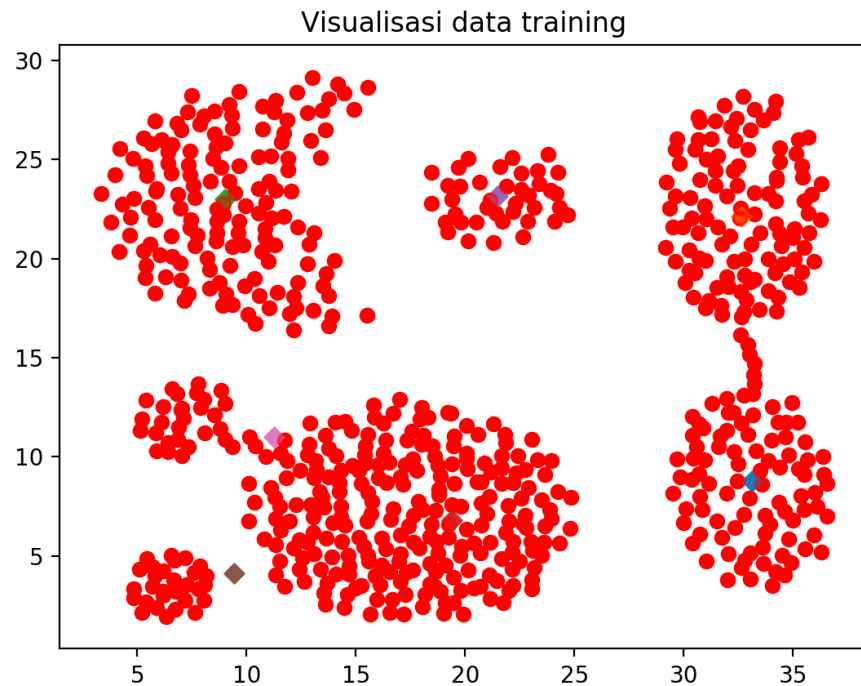


Sarah Fauziah Lestari

1301154552

Teknik Informatika
Fakultas Informatika
Telkom University
2017

A. Pembangunan Model



Dalam tahap awal Training set digunakan sebagai input. Training set berjumlah 600 data, berikut adalah tahap pembangunan model untuk mencari cluster beserta centroid terbaik:

1. Memasukkan file txt kedalam list
2. Menentukan K(cluster) awal saya mencoba dari 2 sampai 12
3. Penentuan centroid secara acak berdasarkan distribusi uniform dikarena dalam visulisasi Training set (gambar 1) data berada pada persebaran $[[0,40],[0,35]]$
4. Setelah mendapatkan centroid dan k maka dilakukan lan perhitungan jarak dengan menggunakan Euclidean
5. Jarak terdekat antara data dan seluruh centroid dihitung dan kemudian diambil sebuah nilai minimum (jarak terdekat data dengan centroid)
6. Lakukan tahap 5 sebanyak cluster yang ada untuk mengetahui cluster berapa mana data tersebut berada.
7. Setelah menemukan cluster untuk masing-masing data seluruh data di cluster masing-masing dicari rata-rata untuk menentukan centroid baru
8. Lakukan perulangan diatas dengan mengganti centroid baru hingga centroid baru tersebut konvergen(stabil), centroid awal = centroid baru

Setelah itu didapatkan centroid yang optimum dengan $k = 7$ dengan nilai :

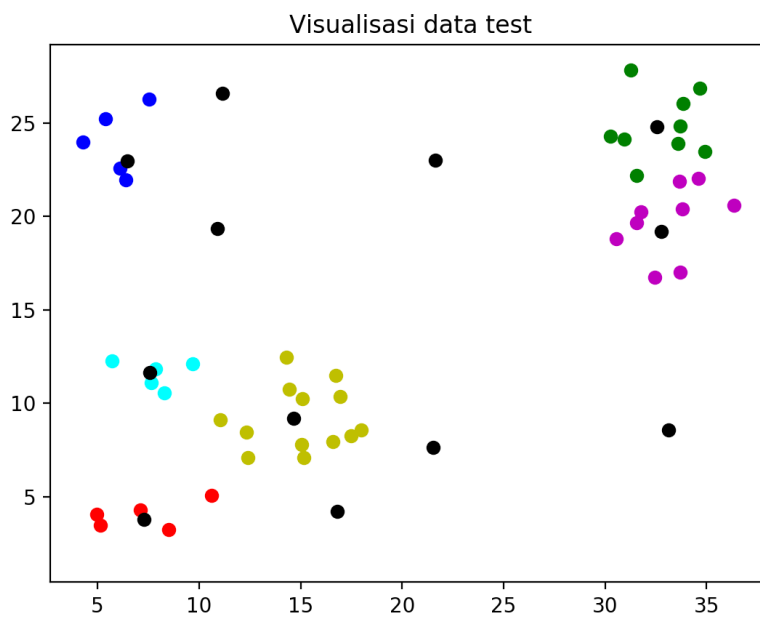
[33.10967741935484, 8.782258064516133], [32.65272727272726,
22.114545454545446], [9.017266187050357, 22.982374100719433],
[19.39664634146342, 6.800304878048779], [21.471249999999999, 23.16875],
[9.44473684210526, 4.123684210526316], [11.273529411764706,
10.991176470588242]]

B. Desain

Setelah di dapat k dan centroid yang optimum maka Test set dijadikan sebagai input yang kemudian melalui tahap sebagai berikut :

1. Melakukan perthitungan jarak antara data dengan centroid yang dipilih.
2. Dipilihlah jarak terdekat untuk mendapatkan cluster per-data

Dari hasil tahapan diatas maka didapat visualisasi data, dan prediksi cluster data di test set sebagai berikut :



Prediksi

[Cluster , [x,y]]
5, [18.75, 22.95]]
[5, [21.45, 21.45]]
[5, [20.5, 22.85]]
[5, [20.65, 24.3]]
[5, [21.7, 23.8]]
[5, [23.1, 21.7]]
[3, [13.35, 28.45]]
[3, [12.4, 27.85]]
[3, [12.2, 28.65]]
[3, [12.9, 26.5]]
[3, [11.15, 28.7]]
[3, [10.5, 28.35]]
[3, [10.25, 27.25]]
[3, [12.6, 24.05]]
[3, [10.05, 25.95]]
[3, [8.5, 27.05]]
[3, [7.55, 26.3]]
[3, [9.4, 25.55]]
[3, [10.55, 24.35]]
[3, [5.4, 25.25]]
[3, [4.3, 24.0]]
[3, [6.1, 22.6]]
[3, [6.4, 21.95]]
[3, [8.45, 17.2]]
[3, [12.3, 22.75]]
[3, [9.95, 19.8]]
[3, [12.0, 20.0]]
[3, [11.4, 19.25]]
[3, [15.2, 18.2]]
[1, [31.9, 4.4]]
[1, [32.8, 6.0]]
[1, [35.45, 4.1]]
[1, [32.3, 7.65]]
[1, [31.8, 9.45]]

[1, [34.7, 8.0]]
[1, [34.55, 8.85]]
[1, [31.2, 12.0]]
[1, [33.1, 12.75]]
[1, [35.8, 11.55]]
[1, [34.05, 13.05]]
[7, [11.05, 9.1]]
[7, [14.3, 12.45]]
[7, [14.45, 10.75]]
[7, [12.35, 8.45]]
[7, [12.4, 7.1]]
[7, [15.1, 10.25]]
[4, [16.75, 11.5]]
[4, [15.05, 7.8]]
[6, [10.6, 5.05]]
[6, [13.9, 5.3]]
[6, [12.75, 3.0]]
[4, [15.15, 7.1]]
[4, [15.25, 2.7]]
[4, [15.7, 6.35]]
[4, [17.3, 4.8]]
[4, [18.15, 1.95]]
[4, [20.45, 2.8]]
[4, [18.65, 4.75]]
[4, [16.6, 7.95]]
[4, [20.8, 4.7]]
[4, [20.55, 5.75]]
[4, [17.5, 8.25]]
[4, [18.0, 8.55]]
[4, [20.05, 6.95]]
[4, [24.15, 4.55]]
[4, [20.0, 8.2]]
[4, [16.95, 10.35]]
[4, [19.0, 11.65]]
[4, [20.7, 10.65]]
[4, [20.95, 10.2]]
[4, [21.75, 8.2]]

[4, [23.0, 7.35]]
[4, [23.7, 8.85]]
[6, [5.15, 3.45]]
[6, [4.95, 4.05]]
[6, [7.1, 4.3]]
[6, [8.5, 3.25]]
[2, [32.45, 16.75]]
[2, [30.55, 18.8]]
[2, [31.55, 19.65]]
[2, [33.7, 17.0]]
[2, [31.75, 20.25]]
[2, [31.55, 22.2]]
[2, [30.95, 24.15]]
[2, [33.65, 21.9]]
[2, [33.8, 20.4]]
[2, [36.35, 20.6]]
[2, [34.6, 22.05]]
[2, [34.9, 23.5]]
[2, [33.6, 23.9]]
[2, [33.7, 24.85]]
[2, [30.25, 24.3]]
[2, [31.25, 27.85]]
[2, [33.85, 26.05]]
[2, [34.65, 26.85]]
[7, [9.7, 12.1]]
[7, [5.7, 12.25]]
[7, [7.85, 11.85]]
[7, [7.65, 11.1]]
[7, [8.3, 10.55]]