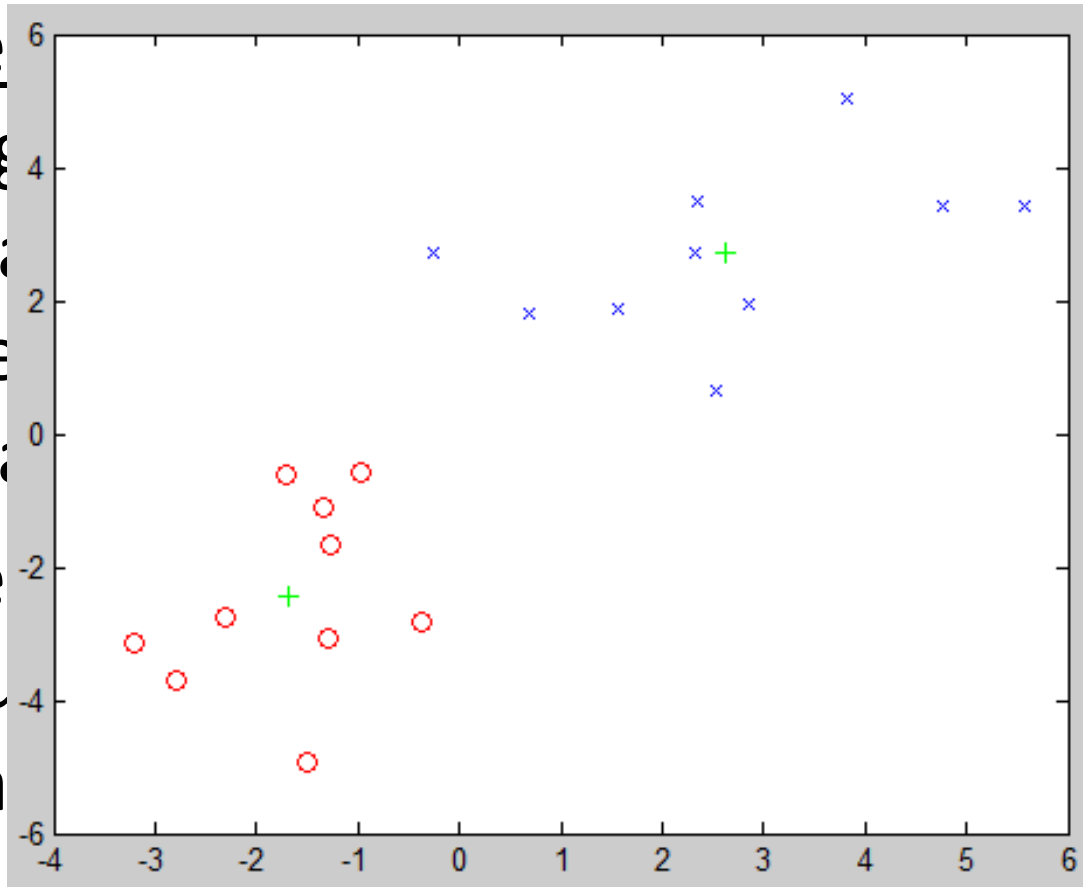


CLUSTERING (K-MEANS)

PENGANTAR

- Clustering mining beberapa segmen beberapa
- Clustering sebelum para a



ta
e dalam
r atau
mpati
ahuan
juga dari
a.

PENGANTAR

- Data tentang 8 nasabah yang pernah memperoleh kredit dari BNI

NASABAH	JUMLAH RUMAH	JUMLAH MOBIL
A	1	3
B	3	3
C	4	3
D	5	3
E	1	2
F	4	2
G	1	1
H	2	1

PROTOTYPE MASALAH PENGGELOMPOKAN

KELOMPOK (CLUSTER)	ANGGOTA KELOMPOK
1	{B}
2	{A,E,G,H}
3	{C,D,F}

- KELOMPOK PERTAMA: KELOMPOK UNIK KARENA MEMILIKI 1 ORG ANGGOTA → (3 RUMAH DAN 3 MOBIL)
- KELOMPOK KEDUA: 4 ORANG ANGGOTA → MEMILIKI RATA-RATA JUMLAH RUMAH SEDIKITNYA 1,25 BUAH DAN RATA-RATA JUMLAH MOBIL 1,75
- KELOMPOK KETIGA: 3 ORANG ANGGOTA → MEMILIKI RATA-RATA JUMLAH RUMAH SEDIKITNYA 4,33 BUAH DAN MOBIL 2,67 BUAH

LANGKAH ALGORITMA K-MEANS

- **LANGKAH PERTAMA:** TANYAKAN PADA PEMAKAI ALGORITMA K-MEANS, BERAPA KELOMPOK YANG INGIN DIBUAT!
- **LANGKAH KEDUA:** SECARA SEMBARANG, PILIH K BUAH CATATAN (DARI SEKIAN CATATAN YANG ADA) SEBAGAI PUSAT-PUSAT KELOMPOK AWAL
- **LANGKAH KETIGA:** UNTUK SETIAP CATATAN, TENTUKAN PUSAT KELOMPOK TERDEKAT DAN TETAPKAN CATATAN TERSEBUT SEBAGAI ANGGOTA DARI KELOMPOK YANG TERDEKAT PUSAT KELOMPOKNYA.
 - HITUNG RASIO BETWEEN CLUSTER VARIATION DENGAN WITHIN CLUSTER VARIATION.
 - BANDINGKAN DENGAN RASIO SEBELUMNYA, JIKA MEMBESAR LANJUT KE LANGKAH KEEMPAT, JIKA TIDAK HENTIKAN PROSES
- **LANGKAH KEEMPAT:** PERBAHARUI PUSAT-PUSAT KELOMPOK (BERDASARKAN LANGKAH KETIGA) DAN KEMBALILAH KE LANGKAH KETIGA

LANGKAH PERTAMA

- APA YANG MENJADI LANGKAH PERTAMA?
 - MENANYAKAN BERAPA KELOMPOK YANG INGIN DIBUAT
 - JIKA KELOMPOKNYA TIGA, MAKA K-NYA ADALAH 3
ATAU $K = 3$

LANGKAH KEDUA

- PILIH SECARA SEMBARANG, PUSAT-PUSAT KELOMPOK ($K=3$)
 - CATATAN B SEBAGAI PUSAT KELOMPOK 1
SEHINGGA $M1=(3,3)$
 - CATATAN E SEBAGAI PUSAT KELOMPOK 2
SEHINGGA $M2=(1,2)$
 - CATATAN F SEBAGAI PUSAT KELOMPOK 3
SEHINGGA $M3=(4,2)$

LANGKAH KETIGA

- TENTUKAN PUSAT KELOMPOK TERDEKATNYA

CATATAN	JARAK KE M1 (3,3)	JARAK KE M2 (1,2)	JARAK KE M3 (4,2)	JARAK TERDEKAT KE KELOMPOK
A (1,3)	2	1	3,162	
B (3,3)	0	2,236	1,414	
C (4,3)	1	3,162	1	C3
D (5,3)	2	4,123	1,414	C3
E (1,2)	2,236	0	3	C2
F (4,2)	1,414	3	0	C3
G (1,1)	2,828	1	3,162	C2
H (2,1)	2,236	1,414	2,236	C2

$$\sqrt{(1-4)^2 + (3-2)^2}$$

LANJUTAN LANGKAH 3 ...

- KELOMPOK YANG DIPEROLEH:
 - KELOMPOK 1 (C1) = {B}
 - KELOMPOK 2 (C2) = {A,E,G,H}
 - KELOMPOK 3 (C3) = {C,D,F}
- HITUNG RASIO BCV DAN WCV
 - $BCV = d(M1,M2) + d(M1,M3) + d(M2,M3) = \underline{6,650}$
 - $WCV = 1^2 + 0^2 + 1^2 + 1,414^2 + 0^2 + 0^2 + 1^2 + 1,414^2$
 $= \underline{7}$
 - $BCV/WCV = 0,950$

LANGKAH KEEMPAT

- TENTUKAN PUSAT KELOMPOK YANG TELAH DITEMUKAN
 - $M1 = \text{RATA-RATA (B)} = (3,3)$
 - $M2 = \text{RATA-RATA (A, E, G, H)} = (1,25; 1,75)$]
 - $M3 = \text{RATA-RATA (C, D, F)} = (4,333; 2,667)$
- KEMBALI KE LANGKAH KETIGA...

LANGKAH KETIGA

- TENTUKAN PUSAT KELOMPOK TERDEKATNYA

CATATAN	JARAK KE M1	JARAK KE M2	JARAK KE M3	JARAK TERDEKAT KE KELOMPOK
A	2	1,275	3,350	C2
B	0	1,768	1,374	C1
C	1	3,021	0,471	C3
D	2	3,953	0,745	C3
E	2,236	0,354	3,399	C2
F	1,414	2,813	0,745	C3
G	2,828	0,791	3,727	C2
H	2,236	1,061	2,867	C2

LANJUTAN LANGKAH 3 ...

- KELOMPOK YANG DIPEROLEH:
 - KELOMPOK 1 (C1) = {B}
 - KELOMPOK 2 (C2) = {A,E,G,H}
 - KELOMPOK 3 (C3) = {C,D,F}
- HITUNG RASIO BCV DAN WCV
 - $BCV = d(M1,M2) + d(M1,M3) + d(M2,M3) = \underline{\underline{6,741}}$
 - $WCV = 1,275^2 + 0^2 + 0,471^2 + 0,745^2 + 0,354^2 + 0,745^2 + 0,791^2 + 1,061^2 = \underline{\underline{4,833}}$
 - $BCV/WCV = \underline{\underline{1,394}}$

LANGKAH KEEMPAT

- TENTUKAN PUSAT KELOMPOK YANG TELAH DITEMUKAN
 - $M1 = \text{RATA-RATA (B)} = (3,3)$
 - $M2 = \text{RATA-RATA (A, E, G, H)} = (1,25;1,75)$
 - $M3 = \text{RATA-RATA (C, D, F)} = (4,333; 2,667)$
- KEMBALI KE LANGKAH KETIGA...

LANGKAH KETIGA

- TENTUKAN PUSAT KELOMPOK TERDEKATNYA

CATATAN	JARAK KE M1	JARAK KE M2	JARAK KE M3	JARAK TERDEKAT KE KELOMPOK
A	2	1,275	3,350	C2
B	0	1,768	1,374	C1
C	1	3,021	0,471	C3
D	2	3,953	0,745	C3
E	2,236	0,354	3,399	C2
F	1,414	2,813	0,745	C3
G	2,828	0,791	3,727	C2
H	2,236	1,061	2,867	C2

LANJUTAN LANGKAH 3 ...

- KELOMPOK YANG DIPEROLEH:
 - KELOMPOK 1 (C1) = {B}
 - KELOMPOK 2 (C2) = {A,E,G,H}
 - KELOMPOK 3 (C3) = {C,D,F}
- HITUNG RASIO BCV DAN WCV
 - $BCV = d(M1,M2) + d(M1,M3) + d(M2,M3) = \underline{\underline{6,741}}$
 - $WCV = 1,275^2 + 0^2 + 0,471^2 + 0,745^2 + 0,354^2 + 0,745^2 + 0,791^2 + 1,061^2 = \underline{\underline{4,833}}$
 - $BCV/WCV = \underline{\underline{1,394}}$

ENDING

- KARENA RASIO BCV/WCV TERAKHIR SAMA DENGAN YANG SEBELUMNYA YAITU 1,394
 - ALGORITMA DIHENTIKAN
 - 3 KELOMPOK YANG DIPEROLEH YAITU
 - KELOMPOK 1 (C_1) = {B}
 - KELOMPOK 2 (C_2) = {A,E,G,H}
 - KELOMPOK 3 (C_3) = {C,D,F}

CONTOH SOAL

- DENGAN MENGGUNAKAN ALGORITMA K-MEANS, CARILAH PENGELOMPOKAN TERBAIK SEHINGGA MEMPEROLEH 2 KLASSTER DARI DATA-DATA BERIKUT:
- $A=(2, 5.0)$, $B=(2, 5.5)$, $C=(5, 3.5)$, $D=(6.5, 2.2)$,
 $E=(7, 3.3)$, $F=(3.5, 4.8)$, $G=(4, 4.5)$

JAWABAN

- $\{A, B, F, G\}$ anggota C_1 dan $\{C, D, E\}$ anggota C_2
- K-means

procedure

```
[coba,pusat]= kmeans (b,2,'dist','sqEuclidean');
```