

Praktikum Membuat Aplikasi AI Unsupervised Learning dengan Algoritma K Means

Oleh :

Harry Witriyono, M.Kom
NIDN. 0210126903

Program Studi Teknik Informatika, Fakultas Teknik
Universitas Muhammadiyah Bengkulu

Apa itu Unsupervised Learning?

- Unsupervised Learning adalah metode pembelajaran mesin di mana **model belajar dari data yang tidak diberi label**. Artinya, sistem mencoba **menemukan pola atau struktur tersembunyi** tanpa mengetahui jawaban yang benar.
- Contoh:
 - Mengelompokkan pelanggan berdasarkan perilaku belanja (clustering)
 - Mengidentifikasi segmentasi gambar tanpa label (image segmentation)
 - Deteksi anomali (anomaly detection)

Perbandingan: Supervised vs. Unsupervised Learning

Aspek	Supervised Learning	Unsupervised Learning
Data	Berlabel (ada target/kelas)	Tidak berlabel
Tujuan	Prediksi kelas atau nilai	Temukan pola, struktur, atau cluster
Contoh Algoritma	KNN, SVM, Decision Tree, Regression	K-Means, DBSCAN, PCA
Contoh Kasus	Memprediksi jenis bunga Iris	Mengelompokkan warna secara alami
Kelebihan	Akurat jika data lengkap	Dapat menemukan pola tak terduga
Kekurangan	Perlu data berlabel (manual)	Hasil bisa ambigu/bervariasi

Mengapa Menggunakan K-Means?

- **Sederhana & Cepat:** Cocok untuk dataset kecil–menengah
- **Intuitif:** Berdasarkan jarak titik ke pusat cluster
- **Tidak perlu label:** Langsung bisa mengelompokkan tanpa pelatihan awal.

Bagaimana Cara Kerja K-Means ?

1. Tentukan jumlah cluster (K)
2. Pilih K titik pusat (Centroid) awal secara random.
3. Hitung jarak setiap data ke masing-masing centroid
4. Kelompokkan ke cluster terdekat.
5. Hitung ulang centroid sehingga stabil (konvergen) Cluster yang dibutuhkananya (dalam contoh data Merah / Biru)

Praktikum Pada Spreadsheet

Untuk data file pembelajaran silahkan lihat di GitHub saya :

<https://github.com/HarryWitriyono/knnwarna2025>

Download file Excel pada folder KMeansWarna

1. Tentukan Jumlah Kluster / Kelas

- Pada contoh ini kita akan menentukan sebanyak 2 kelas yaitu Merah dan Biru.

	A	B	C	D	E	
1	Tabel Data :					
2	No	Kecerahan	Kejenuhan	Kelas		
3	1	40	20		C1	
4	2	50	50		C2	
5	3	60	90			
6	4	10	25			
7	5	70	70			
8	6	60	10			
9	7	25	80			
10					C1	C2
11	Jumlah Kelas yang diharapkan :			2	Merah	Biru

2. Tentukan Titik Awal Kelas

- Penentuan awal ini secara random / acak, pada contoh kita ambil data nomor 1 dan nomor 2 sesuai dengan rencana kita tadi akan membagi sebanyak 2 kelas / Cluster.

	A	B	C	D	E	F
1	Tabel Data :					
2	No	Kecerahan	Kejenuhan	Kelas		
3	1	40	20		C1	
4	2	50	50		C2	
5	3	60	90			
6	4	10	25			
7	5	70	70			
8	6	60	10			
9	7	25	80			
10					C1	C2
11	Jumlah Kelas yang diharapkan :			2	Merah	Biru

3. Hitung Jarak semua data dengan Jarak Awal

- Rumus jaraknya menggunakan perhitungan jarak data Euclidean, yaitu : Akar dari Jumlah dari Kuadrat Selisih masing-masing parameter antara data centroid dengan data yang ada.
- Contoh rumus pada aplikasi Spreadsheet baik pada Google Sheet ataupun Microsoft Excel : `=SQRT(POWER(B17-B3;2)+POWER(C17-C3;2))`. Rumus tersebut menggunakan fungsi SQRT untuk menghitung Akar, dan POWER untuk menghitung Pangkat. Pada Contoh tersebut ada 2 parameter, dan data Centroid berada di Sel B3 untuk parameter 1 dan C3 untuk parameter 2.

4. Tentukan Kelas Iterasi dan Tampilan Pada Spreadsheet

- Sel D17 dibuat rumus :

$$=SQRT(POWER(B17- \$B\$3;2)+POWER(C17- \$C\$3;2))$$
- Sel E17 dibuat rumus :

$$=SQRT(POWER(B17- \$B\$4;2)+POWER(C17- \$C\$4;2))$$
- Untuk menghitung Kelas Iterasinya : jika nilainya jarak Centroid 1 < Centroid 2 maka Kelasnya adalah Kelas Centroid 1, jika tidak Kelas pada Centroid 2. Rumus pada sel F17 :

$$=IF(D17<E17; \$E\$11; \$F\$11)$$

	A	B	C	D	E	F
2	No	Kecerahan	Kejenuhan	Kelas		
3	1	40	20		C1	
4	2	50	50		C2	
5	3	60	90			
6	4	10	25			
7	5	70	70			
8	6	60	10			
9	7	25	80			
10					C1	C2
11	Jumlah Kelas yang diharapkan :			2	Merah	Biru
12						
13	Iterasi	1				
14	Menentukan Centroid Awal secara acak, saya ambil data nomor 1 dan data nom					
15	Jarak Iterasi		1			
16	No	Kecerahan	Kejenuhan	Jarak ke C1	Jarak Ke C2	Kelas Iterasi 1
17	1	40	20	0	31,6227766	Merah
18	2	50	50	31,6227766	0	Biru
19	3	60	90	72,8010989	41,2310563	Biru
20	4	10	25	30,4138127	47,1699057	Merah
21	5	70	70	58,3095189	28,2842712	Biru
22	6	60	10	22,3606798	41,2310563	Merah
23	7	25	80	61,8465844	39,0512484	Biru

5. Hitung Centroid Baru

- Centroid Baru dari hasil Iterasi 1 adalah rata-rata jarak pada masing-masing parameter untuk yang kelasnya sama.
- Pada sel A2 Centroid 1 untuk Parameter 1, rumusnya :
=AVERAGEIF(F17:F23;F17;B17:B23)
- Pada sel B2 :
=AVERAGEIF(F17:F23;F17;C17:C23)
- Untuk Centroid 2 Parameter 1, rumus pada sel A28 :
=AVERAGEIF(F17:F23;F18;B17:B23)
- Untuk Centroid 2 Parameter 2, rumus pada sel B28 :
=AVERAGEIF(F17:F23;F18;C17:C23)

	A	B	C	D	E	F
25	Rata-Rata Jarak Per Kelas Iterasi 1 sebagai centroid baru pada iterasi 2					
26	Kecerahan	Kejenuhan				
27	36,6666667	18,333333	C1			
28	51,25	72,5	C2			
29						
30	Jarak Iterasi		2			
31	No	Kecerahan	Kejenuhan	Jarak ke C1	Jarak Ke C2	Kelas Iterasi 2
32	1	40	20	3,72677996	53,691829	Merah
33	2	50	50	34,3592135	22,5346955	Biru
34	3	60	90	75,3694604	19,5655948	Biru
35	4	10	25	27,4873708	62,9111477	Merah
36	5	70	70	61,4862225	18,9159324	Biru
37	6	60	10	24,7767812	63,1095278	Merah
38	7	25	80	62,7605679	27,3004121	Biru
39						

6. Ulang perhitungan Jarak, Kelas, dan Bandingkan Hasil Kelas Iterasinya

- Ulangi perhitungan jarak setiap data dengan Centroid hasil iterasi sebelumnya, tentukan kelasnya Kembali dan bandingkan dengan hasil kelas pada iterasi sebelumnya. Jika tidak ada perubahan kelas dari iterasi sebelumnya berarti sudah Konvergen atau sudah stabil pembentukan kelasnya.

7. Hasil Akhir

- Hasil akhir penentuan kelas dari semua data yang ada didapat / selesai perulangan perhitungannya jika semua kelas hasil iterasinya tidak berubah posisinya dibandingkan dengan iterasi sebelumnya.
- Dengan kata lain iterasi berhenti saat konvergen / stabil dan tidak berubah lagi posisi Kelasnya pada clusterisasi data yang ada.

	A	B	C	D	E	F	G	H	I
26	Kecerahan	Kejenuhan							
27	36,6666667	18,3333333	C1						
28	51,25	72,5	C2						
29									
30	Jarak Iterasi		2						
31	No	Kecerahan	Kejenuhan	Jarak ke C1	Jarak Ke C2	Kelas Iterasi 2			
32	1	40	20	3,72677996	53,691829	Merah			
33	2	50	50	34,3592135	22,5346955	Biru			
34	3	60	90	75,3694604	19,5655948	Biru			
35	4	10	25	27,4873708	62,9111477	Merah			
36	5	70	70	61,4862225	18,9159324	Biru			
37	6	60	10	24,7767812	63,1095278	Merah			
38	7	25	80	62,7605679	27,3004121	Biru			
39									
40	Karena iterasi 2 sudah tidak berubah dari iterasi sebelumnya, maka hasil Perhitungan Cluster yang didapat :								
41	No	Kecerahan	Kejenuhan	Kelas					
42	1	40	20	Biru					
43	2	50	50	Biru					
44	3	60	90	Merah					
45	4	10	25	Merah					
46	5	70	70	Biru					
47	6	60	10	Biru					
48	7	25	80	Merah					

Tugas

- Buat Aplikasi CRUD dan Perhitungan K-Means untuk menentukan kelas warna berdasarkan nilai Kejenuhan dan Kecerahan. Gunakan database MySQL KNNWarna lalu dan coba bandingkan hasil K-Means dengan kenyataan Data Training yang ada. Buat dalam bahasa pemrograman Web (HTML, CSS, JavaScript, PHP murni OOP).