

APLIKASI K-MEANS UNTUK MENGELOMPOKKAN MAHASISWA BERDASARKAN STATUS GIZI DAN TEKANAN DARAH

Sahat Sonang S¹, Arifin Tua Purba² dan Ferri Ojak Imanuel Pardede³

¹Manajemen Informatika, AMIK Ternas Bangsu

²Teknik Informatika, AMIK Parhina Nusantara

Email : sahatsonangstg@gmail.com, arifintuapuri@gmail.com, ferri_pardede@yahoo.com

ABSTRAKSI

Kesehatan merupakan hal yang sangat perlu untuk diperhatikan, diantaranya mengenai status gizi dan tekanan darah. Dengan mengetahui status gizinya, maka orang tersebut dapat mengontrol berat badan sehingga dapat mencapai berat badan normal yang sesuai dengan tinggi badan. Sedangkan apabila orang tersebut dapat mengetahui tekanan darahnya, maka orang tersebut dapat mengontrol pola makanannya agar selalu berada pada keadaan normal.

Pada penelitian ini, penulis mencoba membangun suatu sistem untuk mengelompokkan data yang ada berdasarkan status gizi dan tekanan darah dengan memasukkan parameter kondisi fisik dari orang tersebut. Pengelompokan data dilakukan dengan menggunakan metode clustering K-Means yaitu dengan mengelompokkan n buah objek ke dalam k kelas berdasarkan jaraknya dengan pusat kelas.

Dari hasil penelitian terhadap 20 data sampel diperoleh 2 kelompok mahasiswa berdasarkan status gizi dan tekanan darah, yaitu : status gizi kurang dan tekanan darah normal, status gizi normal dan tekanan darah normal, status gizi normal dan tekanan darah normal tinggi.

Kata kunci : Clustering, Status Gizi, Tekanan Darah, K-Means.

1. PENDAHULUAN

1.1. Latar Belakang

Seiring perkembangan jaman peran komputer semakin mendominasi kehidupan. Lebih dari itu, komputer saat ini diharapkan dapat digunakan untuk mengerjakan segala sesuatu yang bisa dikerjakan oleh manusia baik di rumah tangga, industri bahkan di lingkungan pendidikan.

Untuk memecahkan masalah dengan komputer, program harus dibuat terlebih dahulu kemudian akan diproses selanjutnya. Tanpa program, komputer hanyalah sebuah kotak besi yang tidak berguna. Masalah penentuan status gizi dan tekanan darah merupakan hal yang sering terlupakan oleh mahasiswa yang pada umumnya selalu disibukkan dengan

berbagai kegiatan sehari-hari. Sering ditemui seorang mahasiswa tidak mengetahui berada di kelompok mana untuk setiap kriteria diatas. Dengan mengetahui kelompoknya maka mahasiswa dapat mengambil tindakan agar selalu berada dalam kelompok normal.

K-Means merupakan salah satu metode data clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster/kelompok. Metode ini mempartisi data ke dalam cluster/kelompok sehingga data yang memiliki karakteristik sama dikelompokkan ke dalam satu cluster yang sama.

1.2. Tujuan

Tujuan penelitian ini adalah membangun aplikasi untuk mengelompokkan mahasiswa berdasarkan status gizi, dan tekanan darah menggunakan metode klasifikasi K-Means.

2. DASAR TEORI

2.1. Algoritma Klasifikasi K-Means

Algoritma K-means adalah implementasi dari algoritma clustering partisional yang paling sederhana dan paling banyak digunakan. K-means juga menggunakan kuadrat error criterion. Algoritma ini mulai mempartisi ruang data secara acak sambil penunjukan (assignment) sampel yang ada ke dalam kluster-kluster berdasarkan kemiripan antara kluster dan sampel, sampai sebuah criterion yang convergen ditemukan. Syarat sebuah criterion telah ditemukan adalah ketika tidak ada lagi pemindahan sampel (reassignment) dari satu kluster ke kluster yang lain yang akan menyebabkan berkurangnya total error yang dikuadratkan (errorsquare). Algoritma ini populer digunakan karena kemudahan implementasinya, dan memiliki kecepatan yang cukup baik.

Langkah-langkah dasar yang diambil oleh algoritma

K-means adalah:

- Langkah pertama: Tanyakan kepada pemakai algoritma K-means, catatan-catatan yang ada akan dibuat menjadi beberapa kelompok, sebutlah sebanyak k kelompok.
- 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 terdekatnya dan letakkan catatan tersebut sebagai anggota dari kelompok yang terdekat pusat kelompoknya. Hitung rasio antara besaran Between Cluster Variation dengan Within Cluster Variation, lalu bandingkan rasio tersebut dengan rasio sebelumnya (bila sudah ada). Jika rasio tersebut membesar, maka lanjutkan ke langkah keempat. Jika tidak hentikan prosesnya.
- Langkah keempat: perbaharui pusat-pusat kelompok (berdasarkan kelompok yang didapat dari langkah ketiga) dan kembali ke langkah ketiga.

2.2. Status Gizi

Status gizi adalah salah satu indikator kesehatan masyarakat yang amat penting untuk dievaluasi secara periodik. Kelebihan gizi atau status gizi yang lebih dapat berdampak buruk terhadap kesehatan seseorang seperti halnya dengan obesitas. Obesitas merupakan suatu kondisi yang kronis dengan karakteristik kelebihan lemak tubuh dan hal itu sekarang merupakan masalah medik yang prevalensinya meningkat setiap waktu. Menurut laporan FAO/WHO/UNU tahun 1985 menyatakan bahwa batasan berat badan orang dewasa ditentukan berdasarkan nilai Body Mass Index diterjemahkan menjadi Index Massa Tubuh (IMT). Index Massa Tubuh merupakan alat yang sederhana untuk menentukan status gizi orang dewasa khususnya yang berkaitan dengan kekurangan dan kelebihan berat badan. Dalam menghitung IMT diperlukan dua parameter, yaitu berat badan (cm) dan tinggi badan (m). IMT dapat dihitung dengan menggunakan persamaan,

$$IMT = \frac{\text{berat badan (kg)}}{(\text{tinggi badan (cm)})^2} \dots \dots \dots (1)$$

Dengan batas pengelompokan pada Tabel 1

Tabel 1. Batas Pengelompokan

Klasifikasi	IMT (kg/m ²)	Risiko Penyakit
Kurang	< 18,5	Kemah (tetapi resiko problem klinik lain meningkat)
Normal	18,5 - 24,9	Sesuai
Berat badan lebih	25,0 - 29,9	Meningkat
Obes Kelas I	30,0 - 34,9	Sedang
Obes Kelas II	35,0 - 39,9	Berat
Obes Kelas III	≥ 40,0	Sangat berat

2.3. Tekanan Darah

Hipertensi atau darah tinggi adalah keadaan dimana seseorang mengalami peningkatan tekanan darah diatas norma. Hipertensi merupakan kelainan yang sulit diketahui oleh tubuh kita sendiri. Satu-satunya cara untuk mengetahui hipertensi adalah dengan mengukur tekanan darah kita secara teratur.

Secara umum seseorang dikatakan menderita hipertensi jika tekanan darah sistolik/diastoliknya > 140/90 mmHg atau mengkonsumsi obat antihipertensi. Sistolik adalah tekanan darah pada saat jantung memompa darah ke dalam pembuluh nadi (saat jantung mengkerut). Diastolik adalah tekanan darah pada saat jantung mengembang dan menyedot darah kembali (pembuluh darah mengempis kosong).

Berdasarkan hasil pertemuan ketujuh Komite Nasional Gabungan Amerika Serikat untuk prevensi, deteksi, evaluasi, dan pengobatan tekanan darah tinggi, hipertensi didefinisikan bila tekanan darah sistolik mencapai 140 mmHg atau lebih, diastolik mencapai 90 mmHg atau lebih. Berikut ini merupakan klasifikasi tekanan darah menurut The Seventh Report of The Joint National Committee on Prevention, Detection, Evaluation, and Treatment of high Blood Pressure (JNC, 2001).

Tabel 2. Klasifikasi Tekanan Darah

Kategori	Sistolik (mmHg)	Diastolik (mmHg)
Normal	< 130	< 85
Normal Tinggi	130 - 139	85 - 89
Hipertensi Ringan	140 - 159	90 - 99
Hipertensi Sedang	160 - 179	100 - 109
Hipertensi Berat	180 - 209	110 - 119
Hipertensi Maliga	≥ 210	≥ 120

3. PENGUJIAN

3.1. Tekanan Darah

Perancangan data dengan K-Means Clustering dengan menggunakan data status gizi dan tekanan darah dengan mengambil contoh sebanyak 20 record yaitu record pertama sampai dengan record kedua puluh untuk diuji secara manual.

Tabel 3. Data Mahasiswa

Mhs. Ke	BB	TB	Sistolik	Diastolik
1	45	166	110	90
2	45	161	90	80
3	58	157	120	90
4	43	155	110	80
5	55	165	110	80
6	48	160	110	80
7	50	162	110	80
8	48	155	110	90
9	48	152	110	90
10	47	154	110	90
11	49	155	110	90

Mhs. Ke	BB	TB	Sistolik	Diastolik
12	49	153	110	90
13	51	163	100	80
14	38	159	110	90
15	42	158	110	80
16	46	150	110	80
17	40	150	110	90
18	42	157	110	90
19	52	157	100	90
20	50	158	100	90

Keterangan

BB : Berat Badan

TB : Tinggi Badan

Data pada Tabel 3 kemudian digunakan untuk menghitung nilai IMT dengan menggunakan persamaan (1). Hasil dari perhitungan dapat dilihat pada Tabel 4.

Tabel 4. Hasil Perhitungan IMT dari Tekan

Mhs. Ke	IMT	Sistolik	Diastolik
1	16,73	110	90
2	17,36	90	80
3	23,53	120	90
4	17,90	110	80
5	20,20	110	80
6	18,75	110	80
7	19,05	110	80
8	19,98	110	90
9	20,78	110	90
10	19,83	110	90
11	20,40	110	90
12	20,40	110	90
13	18,75	100	80
14	15,93	110	90
15	19,23	110	80
16	20,44	110	80
17	17,78	110	90
18	17,04	110	90
19	21,10	100	90
20	20,03	100	90

Keterangan

IMT : Index Massa Tubuh

3.2. Hasil Penentuan

Selanjutnya akan digunakan algoritma

klasifikasi K-Means untuk mengelompokkan data yang ada.

- Langkah pertama dari K-Means clustering adalah menanyakan kepada pemakai K-Means, record-record yang ada akan dibuat menjadi berapa kelompok. Jika jumlah kelompoknya tiga, maka nilai k-nya adalah 3 atau k=3.
- Pada langkah kedua ini, kita akan secara sembarang memilih k=3 buah record (dari 20 reco yang ada) sebagai pusat-pusat kelompok awal, misalnya
 - Record ke-2 sebagai Kelompok 1 sehingga C1={27,36,90,80}
 - Record ke-5 sebagai Kelompok 2 sehingga C2={20,20,110,80}
 - Record ke-12 sebagai Kelompok 3 sehingga C3={20,40,110,90}
- Pada langkah ketiga ini, setiap record akan ditentukan pusat kelompok terdekatnya. Record tersebut akan diterapkan sebagai kelompok yang terdekat pusat kelompoknya. Hitung jarak setiap

data yang ada terhadap setiap pusat cluster dengan persamaan berikut:

$$D_{1j} = \sqrt{\sum_{i=1}^3 (x_i - c_{1i})^2}$$

$$D_{11} = \sqrt{(16,23 - 27,36)^2 + (110 - 90)^2 + (90 - 80)^2}$$

$$D_{11} = 24,94$$

$$D_{12} = \sqrt{(16,23 - 23,36)^2 + (110 - 110)^2 + (90 - 80)^2}$$

$$D_{12} = 10,00$$

$$D_{13} = \sqrt{(16,23 - 20,44)^2 + (110 - 90)^2 + (90 - 80)^2}$$

$$D_{13} = 14,07$$

Hasil perhitungan selengkapnya pada Tabel 5.

Tabel 5. Hasil Perhitungan Jarak Setiap Data

No. Data	Jarak	Kategori	Distribusi	Kelompok		
				1	2	3
1	16,23	110	90	24,93	3,67	4,07
2	17,36	90	80	10,90	22,57	22,57
3	23,35	120	90	21,85	10,53	17,48
4	17,90	110	80	22,13	10,76	10,31
5	20,29	110	80	21,24	10,00	10,90
6	18,75	110	80	21,77	10,11	10,14
7	18,65	110	80	21,85	10,07	10,09
8	19,98	110	90	23,55	0,22	0,42
9	20,78	110	90	23,21	0,58	0,38
10	19,85	110	90	23,67	0,38	0,38
11	20,40	110	90	23,42	0,20	0,00
12	20,40	110	90	23,42	0,20	0,00
13	18,73	100	80	13,21	14,22	14,24
14	15,63	110	90	25,73	5,17	5,17
15	19,23	110	80	24,59	10,05	10,07
16	20,44	110	80	21,16	10,00	10,90
17	17,78	110	90	24,43	2,42	2,62
18	17,04	110	90	24,67	3,16	3,30
19	21,10	100	80	15,47	10,04	10,02
20	20,03	100	80	15,97	10,00	10,01

Suatu data akan menjadi anggota dari suatu cluster yang memiliki jarak terkecil dari pusat clusternya. Misalkan untuk data pertama, jarak terkecil diperoleh pada cluster kedua, sehingga data pertama akan menjadi anggota dari cluster kedua. Demikian juga untuk data kedua, jarak terkecil ada pada cluster kepertama, maka data tersebut akan masuk pada cluster kepertama. Posisi cluster selengkapnya dapat dilihat pada Tabel 6.

Tabel 6. Posisi Cluster Pada Iterasi Pertama

No. Data	Jarak	Kategori	Distribusi	Kelompok			JK
				1	2	3	
1	16,23	110	90	24,93	3,67	4,07	C2
2	17,36	90	80	10,90	22,57	22,57	C1
3	23,35	120	90	21,85	10,53	17,48	C3
4	17,90	110	80	22,13	10,76	10,31	C2
5	20,29	110	80	21,24	10,00	10,90	C2
6	18,75	110	80	21,77	10,11	10,14	C2
7	18,65	110	80	21,85	10,07	10,09	C2
8	19,98	110	90	23,55	0,22	0,42	C2
9	20,78	110	90	23,21	0,58	0,38	C2
10	19,85	110	90	23,67	0,38	0,38	C2
11	20,40	110	90	23,42	0,20	0,00	C2
12	20,40	110	90	23,42	0,20	0,00	C2
13	18,73	100	80	13,21	14,22	14,24	C1
14	15,63	110	90	25,73	5,17	5,17	C2
15	19,23	110	80	24,59	10,05	10,07	C2
16	20,44	110	80	21,16	10,00	10,90	C2
17	17,78	110	90	24,43	2,42	2,62	C2
18	17,04	110	90	24,67	3,16	3,30	C2
19	21,10	100	80	15,47	10,04	10,02	C3
20	20,03	100	80	15,97	10,00	10,01	C2

No. Data	Jarak	Kategori	Distribusi	Kelompok			JK
				1	2	3	
11	20,40	110	90	23,42	0,20	0,00	C3
12	20,40	110	90	23,42	0,20	0,00	C1
13	18,73	100	80	13,21	14,22	14,24	C1
14	15,63	110	90	25,73	5,17	5,17	C2
15	19,23	110	80	24,59	10,05	10,07	C2
16	20,44	110	80	21,16	10,00	10,90	C2
17	17,78	110	90	24,43	2,42	2,62	C2
18	17,04	110	90	24,67	3,16	3,30	C2
19	21,10	100	80	15,47	10,04	10,02	C3
20	20,03	100	80	15,97	10,00	10,01	C2

Keterangan

JK = Jarak Terdekat ke Kelompok

Dari Tabel 6 didapat keanggotaan sebagai berikut:

1. Kelompok 1 (atau C1) = { 2,13 }
2. Kelompok 2 (atau C2) = { 1,4,5,6,7,8,10,14,15,17,18,20 }
3. Kelompok 3 (atau C3) = { 3,9,11,12,16,19 }

Pada langkah ini dihitung rasio antara besaran Between Cluster Variation (BCV) dengan Within Cluster Variation (WCV), sebagai berikut:

$$BCV = D(C1,C2) + D(C1,C3) + D(C2,C3)$$

$$BCV = 12,54 + 12,57 + 0,20$$

$$BCV = 25,31$$

$$WCV = \sum_{i=1}^k \sum_{p \in C_i} d(p, m_i)^2$$

$$WCV = 3,87^2 + 10,00^2 + 10,48^2 + 10,26^2 + 10,09^2 + 10,19^2 + 10,07^2 + 0,22^2 + 0,38^2 + 0,38^2 + 0,60^2 + 0,00^2 + 13,21^2 + 5,17^2 + 10,05^2 + 10,00^2 + 2,42^2 + 3,16^2 + 10,02^2 + 10,00^2$$

$$WCV = 1252,29$$

$$BCV/WCV = 25,31/1252,29 = 0,0202$$

belum mendapat rasio ini, maka perbandingan rasio belum dapat dilakukan

dan proses K-Means clustering dilanjutkan ke langkah keempat.

4. Langkah Keempat (Iterasi-2)

Pada langkah ini, pembaharuan pusat-pusat kelompok akan dilakukan seperti berikut:

- a. Untuk kelompok 1, ada 2 data yaitu data ke-2 dan 13, sehingga: $C11 = (17.36 + 18.73)/2 = 18.05$ $C12 = (90 + 100)/2 = 95.00$ $C13 = (80 + 80)/2 = 80.00$ Sehingga $C1 = \{18.05, 95.00, 80.00\}$
- b. Untuk kelompok 2, ada 12 data yaitu data ke-1, 4, 5, 6, 7, 8, 10, 14, 15, 17, 18 dan 20, sehingga: $C21 = (16.33+17.90+20.20+18.75+19.05+19.98+19.82+15.03+19.2+17.78+17.04+20.03)/12 = 18.43$ $C22 = (110+110+110+110+110+110+110+110+110+110+110+100)/12 = 109.17$ $C23 = (90+80+80+80+80+90+90+90+80+90+90+90)/12 = 85.83$ Sehingga $C2 = \{18.43, 109.17, 85.83\}$
- c. Untuk kelompok 3, ada 6 data yaitu data ke-3, 9, 11, 12, 16 dan 19, sehingga: $C31 = (23.53+20.78+20.40+20.40+20.44+21.10)/6 = 21.11$ $C32 = (120+110+110+110+110+100)/6 = 110.00$ $C33 = (90+90+90+90+80+90)/6 = 83.33$ Sehingga $C3 = \{21.11, 110.00, 83.33\}$

5. Ulangi langkah 3 hingga nilai rasio antara besaran Between Cluster Variation (BCV) dengan Within Cluster Variation (WCV) tidak mengalami penakut dari nilai rasio sebelumnya.

Tabel 7. Posisi Cluster Pada Iterasi Kedua

Mhs. Ke	Nilai	Simbol	Distribusi	Kelompok			Jumlah
				1	2	3	
1	16.33	110	90	18.11	4.74	2.00	C3
2	17.36	90	80	5.05	20.06	21.01	C1
3	23.53	120	90	27.48	12.68	10.49	C3
4	17.90	110	80	15.09	5.92	8.93	C2
5	20.20	110	80	15.13	6.15	8.39	C2
6	18.75	110	80	15.02	5.90	8.56	C2
7	19.05	110	80	15.03	5.94	8.55	C2
8	19.98	110	80	15.13	6.52	7.91	C2
9	20.78	110	80	18.23	4.85	1.08	C3
10	19.82	110	80	18.11	4.43	2.31	C3
11	20.40	110	80	18.18	4.68	1.87	C3
12	20.40	110	80	18.18	4.68	1.84	C3
13	18.73	100	80	5.05	19.87	13.25	C1
14	15.03	110	80	15.20	5.44	6.79	C2
15	19.22	110	80	15.03	5.95	8.44	C2
16	20.44	110	80	15.19	6.23	8.26	C2
17	17.78	110	80	18.71	4.50	3.71	C3
18	17.04	110	80	5.06	5.17	4.38	C1
19	21.10	100	80	11.59	19.42	10.14	C3
20	20.03	100	80	11.55	19.26	10.29	C3

$$BCV = D(C1,C2) + D(C1,C3) + D(C2,C3)$$

$$BCV = 15.02 + 2.36 + 3.01$$

$$BCV = 20.39$$

$$WCV = \sum_{i=1}^n \sum_{j=1}^m d(i, m_j)^2$$

$$WCV = 4.74^2 + 5.05^2 + 10.42^2 + 5.92^2 + 6.15^2 + 5.90^2$$

$$+ 5.93^2 + 2.01^2 + 1.70^2 + 7.11^2 + 1.81^2 + 1.81^2$$

$$+ 5.05^2 + 5.44^2 + 5.95^2 + 6.23^2 + 3.73^2 + 4.40^2$$

$$+ 10.14^2 + 10.29^2$$

$$WCV = 686.48$$

Tabel 8. Posisi Cluster Pada Iterasi Ketiga

Mhs. Ke	Nilai	Simbol	Distribusi	Kelompok			Jumlah
				1	2	3	
1	16.33	110	90	18.11	4.74	2.00	C3
2	17.36	90	80	5.05	20.19	21.06	C1
3	23.53	120	90	27.48	12.52	11.33	C3
4	17.90	110	80	15.06	5.94	10.28	C2
5	20.20	110	80	15.13	6.16	10.05	C2
6	18.75	110	80	15.02	5.93	10.34	C2
7	19.05	110	80	15.03	5.99	10.18	C2
8	19.98	110	90	15.13	7.67	1.91	C3
9	20.78	110	80	18.23	7.85	1.22	C3
10	19.82	110	90	18.11	7.64	1.93	C3
11	20.40	110	90	18.18	7.77	1.93	C3
12	20.40	110	90	18.18	7.77	1.93	C3
13	18.73	100	80	5.05	19.31	11.52	C1

Abis Ke	BMI	Sistolik	Diastolik	Kelompok			JK
				1	2	3	
14	15,03	116	90	18,28	8,21	5,15	C3
15	19,23	110	85	15,95	3,64	10,09	C2
16	20,44	110	80	15,19	7,25	10,06	C2
17	27,38	130	90	18,94	7,52	2,51	C3
18	17,04	110	90	18,06	7,62	3,20	C3
19	21,10	100	90	14,59	12,79	9,68	C3
20	20,93	100	90	11,35	12,61	9,09	C3

$$BCV = D(C1,C2) + D(C1,C3) + D(C2,C3)$$

$$BCV = 15,13 + 3,21 + 7,74$$

$$BCV = 26,08$$

$$WCV = \sum_{i=1}^k \sum_{p \in C_i} d(p, m_i)^2$$

$$WCV = 3,88^2 + 5,05^2 + 11,53^2 + 2,54^2 + 3,10^2 + 2,53^2$$

$$+ 2,59^2 + 1,01^2 + 1,22^2 + 1,03^2 + 1,05^2 + 1,05^2$$

$$+ 5,05^2 + 5,15^2 + 2,64^2 + 3,23^2 + 2,51^2 + 3,20^2$$

$$+ 9,06^2 + 9,00^2$$

$$WCV = 457,57$$

BCV/WCV = 26,08/457,57 = 0,057

Tabel 9. Posisi Cluster Pada Iterasi Keempat

Abis Ke	BMI	Sistolik	Diastolik	Kelompok			JK
				1	2	3	
1	18,15	100	90	18,11	10,02	3,33	C3
2	17,26	90	80	5,05	30,09	21,71	C1
3	23,53	120	90	27,43	18,77	11,91	C3
4	17,90	110	80	15,90	1,26	10,14	C2
5	20,29	110	80	15,15	0,94	10,97	C2
6	18,35	100	80	15,02	0,24	10,05	C2
7	19,05	110	80	15,03	0,21	10,04	C2
8	19,99	110	90	18,13	10,05	1,01	C3
9	20,78	110	90	18,73	10,11	1,05	C3
10	19,82	110	90	18,17	10,02	0,98	C3
11	20,40	110	90	17,18	10,06	1,34	C3
12	20,40	110	90	18,18	10,06	1,34	C3
13	19,73	100	80	5,05	10,01	13,28	C1
14	15,94	110	90	18,28	10,86	4,40	C3
15	19,23	110	85	15,95	3,63	10,09	C2
16	20,44	110	80	15,19	7,25	10,06	C2
17	17,78	110	90	18,03	10,11	1,28	C3
18	17,04	110	90	18,06	10,24	2,46	C3
19	21,10	100	90	14,59	12,26	9,33	C3

$$BCV = D(C1,C2) + D(C1,C3) + D(C2,C3)$$

$$BCV = 15,04 + 3,37 + 10,00$$

$$BCV = 28,61$$

$$WCV = \sum_{i=1}^k \sum_{p \in C_i} d(p, m_i)^2$$

$$WCV = 3,13^2 + 5,05^2 + 11,61^2 + 1,36^2 + 0,94^2 + 0,51^2$$

$$+ 0,21^2 + 1,04^2 + 1,65^2 + 0,96^2 + 1,34^2 + 1,34^2$$

$$+ 5,05^2 + 4,40^2 + 0,03 + 1,18^2 + 1,78^2 + 2,46^2$$

$$+ 9,33^2 + 0,19^2$$

$$WCV = 408,46$$

$$BCV/WCV = 28,61/408,46 = 0,0700$$

Tabel 10. Posisi Cluster Pada Iterasi Kelima

Abis Ke	BMI	Sistolik	Diastolik	Kelompok			JK
				1	2	3	
1	18,33	110	90	18,11	10,47	3,13	C3
2	17,36	90	80	5,05	20,09	21,71	C1
3	23,53	120	90	27,43	18,77	11,91	C3
4	17,90	110	80	15,90	1,26	10,14	C2
5	20,29	110	80	15,15	0,94	10,97	C2
6	18,78	110	90	18,02	0,51	10,05	C2
7	19,05	110	80	15,03	0,21	10,04	C2
8	19,99	110	90	18,13	10,03	1,04	C3
9	20,78	110	90	18,24	10,11	1,05	C3
10	19,82	110	90	18,11	10,02	0,98	C3
11	20,40	110	90	18,18	10,06	1,34	C3
12	20,40	110	90	18,18	10,06	1,34	C3
13	18,73	100	80	5,05	10,01	13,28	C1
14	15,03	110	90	18,28	10,86	4,40	C3
15	19,23	110	80	15,05	3,63	10,09	C2
16	20,44	110	80	15,19	7,25	10,06	C2
17	17,78	110	90	18,03	10,11	1,28	C3
18	17,04	110	90	18,06	10,24	2,46	C3
19	21,10	100	90	14,59	11,26	9,33	C3
20	20,93	100	90	11,35	14,10	9,10	C3

$$BCV = D(C1,C2) + D(C1,C3) + D(C2,C3)$$

$$BCV = 15,04 + 3,37 + 10,00$$

$$BCV = 28,61$$

$$WCV = \sum_{i=1}^k \sum_{p \in C_i} d(p, m_i)^2$$

$$WCV = 3,13^2 + 5,05^2 + 11,61^2 + 1,36^2 + 0,94^2 + 0,51^2$$

$$+ 0,21^2 + 1,04^2 + 1,65^2 + 0,96^2 + 1,34^2 + 1,34^2$$

$$+ 5,05^2 + 4,40^2 + 0,03^2 + 1,18^2 + 1,78^2 + 2,46^2$$

$$+ 9,33^2 + 0,19^2$$

$$WCV = 408,46$$

$$BCV/WCV = 28,61/408,46 = 0,0700$$

Karena pada iterasi kelima nilai rasio antara besaran Between Cluster Variation (BCV) dengan Within Cluster Variation (WCV) tidak mengalami pencikaman dari nilai rasio sebelumnya (pada iterasi keempat), maka iterasi di hentikan dan hasil akhir yang diperoleh ada 3 cluster:

- Cluster pertama memiliki pusat (18,05; 95; 80) yang dapat diartikan sebagai kelompok mahasiswa dengan status gizi kurang dan tekanan darah normal (sistolik 95 mmHg dan diastolik 80 mmHg).
- Cluster kedua memiliki pusat (19,26; 110; 80) yang dapat diartikan sebagai kelompok mahasiswa dengan status gizi normal dan tekanan darah normal (sistolik 110 mmHg dan diastolik 80 mmHg).
- Cluster ketiga memiliki pusat (19,35; 109,17; 90) yang dapat diartikan sebagai kelompok mahasiswa dengan status gizi

normal dan tekanan darah normal tinggi (sistolik 109.17 mmHg dan diastolik 90 mmHg).

4. KESIMPULAN

Dari hasil penelitian, dapat disimpulkan bahwa algoritma klasifikasi K-Means dapat digunakan untuk mengelompokkan mahasiswa berdasarkan status gizi dan tekanan darah. Dari data yang dilatih, diperoleh 3 kelompok berdasarkan IMI

dan tekanan darah, yaitu :

- 1) IMT kurang dan tekanan darah normal, dengan pusat cluster (18.05; 95; 80).
- 2) IMT normal dan tekanan darah normal, dengan pusat cluster (19.26; 110; 80).

- 3) IMT normal dan tekanan darah normal tinggi, dengan pusat cluster (19.35; 109.17; 90).

REFERENSI

- [1] Manampiring, E. Aaltje. 2008. Laporan Penelitian "Hubungan Status Gizi dan Tekanan Darah Pada Penduduk Usia 45 Tahun ke Atas di Kelurahan Pakowa Kecamatan Wanca Kota Manado".
- [2] Witten, Ian H. dan Frank, Eibe. 2005. Data Mining Practical Machine Learning Tools and Techniques, Second Edition. Morgan Kaufmann, San Francisco.
- [3] Agusta, Yudi. Pebruari 2007. "K-Means- Penerapan, Permasalahan dan Metode Terkait". Jurnal Sistem dan Informatika Vol.3 : 47-60.
- [4] Susanto, Sani. Dab Suryadi, Dedy. 2010. Pengantar Data Mining. Penerbit Andi Yogyakarta.