

2018-ARS-K-means-Kelly

by Irmeilyana Irmeilyana

Submission date: 27-May-2023 09:24AM (UTC+0700)

Submission ID: 2102862819

File name: 2018-ARS-K-Means-Kelly.pdf (1.09M)

Word count: 4248

Character count: 23125

Analisis Karakteristik Mahasiswa Berdasarkan Nilai Kelompok Mata Kuliah dengan Menggunakan Analisis *Cluster K-Means*

Irmeilyana

Jurusan Matematika
Universitas Sriwijaya
Indralaya, Indonesia
imel_unsri@yahoo.co.id

Kelly Puta Hera

Alumni Jurusan Matematika, Universitas Sriwijaya
Indralaya, Indonesia
kellyputahera@gmail.com

Sugandi Yahdin³

Jurusan Matematika
Universitas Sriwijaya
Indralaya, Indonesia
0819687369a@gmail.com

Rana Sania

Alumni Jurusan Matematika
Universitas Sriwijaya
Indralaya, Indonesia
whereismyatie@gmail.com

1 **Abstrak**—Indeks Prestasi (IP) adalah nilai yang dihitung berdasarkan jumlah beban studi yang diambil dalam satu semester dikalikan dengan bobot prestasi tiap-tiap mata kuliah kemudian dibagi dengan jumlah beban kredit yang diambil. IPK merupakan IP kumulatif dari seluruh semester yang telah diikuti mahasiswa. Pada kurikulum Tahun 2012, Jurusan Matematika FMIPA Unsri mempunyai 5 Kelompok Bidang Minat (KBM), yaitu: Matematika Murni, Optimasi, Statistika, Aktuaria, dan Komputasi. Selain itu ada juga mata kuliah wajib yang terbagi menjadi Mata Kuliah Umum (MKU), Mata Kuliah Dasar (MKD), Mata Kuliah Keilmuan Matematika (MKM), dan Mata Kuliah Pilihan (MKP). Tujuan dari penelitian ini untuk menganalisis karakteristik mahasiswa berdasarkan kelompok mata kuliah dengan menggunakan analisis kluster *K-Means* pada alumni angkatan 2012. Berdasarkan analisis kluster *K-Means* didapatkan 3 kluster atau kelompok mahasiswa. Kluster 1 beranggotakan 10 mahasiswa dan merupakan mahasiswa dengan IPK tinggi serta dominan mengambil KBM Optimasi. Kluster 2 dengan anggota 30 mahasiswa merupakan kelompok mahasiswa yang mempunyai IPK sedang serta dominan mengambil KBM Statistika dan Aktuaria. Selanjutnya kluster 3 yang beranggotakan 17 mahasiswa merupakan kelompok mahasiswa yang memiliki IPK rendah dengan KBM yang dominan diambil adalah KBM Murni. Sementara itu berdasarkan korelasi antar peubah IPK yang tinggi dapat direpresentasikan oleh IP kelompok Mata Kuliah Keilmuan Matematika (MKM) yang juga tinggi.

Kata Kunci :—IPK, Kelompok Mata Kuliah, Kelompok Bidang Minat, *K-Means*.

I. PENDAHULUAN

Indeks Prestasi (IP) adalah nilai yang **1** dihitung berdasarkan jumlah beban studi yang diambil dalam

satu semester dikalikan dengan bobot prestasi tiap-tiap mata kuliah kemudian dibagi dengan jumlah beban kredit yang diambil. IPK merupakan IP kumulatif dari seluruh semester yang telah diikuti mahasiswa (Universitas Sriwijaya, 2013).

3 Pendidikan dikatakan berkualitas bila proses belajar mengajar dapat berjalan dengan lancar, efektif, efisien, dan ada interaksi antara komponen-komponen yang terkandung dalam sistem pengajaran yaitu tujuan pendidikan dan pengajaran, peserta didik, tenaga kependidikan atau guru, kurikulum, strategi pembelajaran, media pengajaran dan evaluasi pengajaran [1].

Pada kurikulum Tahun 2012, Jurusan Matematika FMIPA Unsri mempunyai 5 KBM, yaitu: Matematika Murni, Optimasi, Statistika, Aktuaria, dan Komputasi. Pembagian MK pada kelompok bidang minat ini sangat mendukung visi, misi, dan tujuan Jurusan Matematika, serta tercapainya profil dan kompetensi lulusan Jurusan Matematika yang diharapkan.

Pada angkatan 2011, sebagian besar mahasiswa Jurusan Matematika FMIPA Unsri tidak mempunyai kecenderungan yang lebih menonjol dalam IPK dan IP setiap kelompok MK. Sebagian kecil mahasiswa yang mempunyai IPK yang tinggi ataupun rendah dapat tercermin dari IP Kelompok MK Umum, Dasar, dan IP setiap KBM. Mahasiswa peminat KBM Statistika cenderung mempunyai nilai IPK yang tinggi. IPK mahasiswa yang tinggi selain lebih dicirikan oleh IP Kelompok MK Wajib, juga dicirikan oleh IP Kelompok MK Dasar, Umum, dan IP KBM Statistika [2]. Pada penelitian ini digunakan analisis kluster berhirarki.

Menurut [3], *K-Means* merupakan metode klastering non hirarki yang sederhana dan mudah implementasikan karena memiliki kelebihan dalam mengelompokkan data yang berukuran besar serta menghasilkan klaster yang cepat. Pada *K-Means*, jumlah klaster awal k harus ditentukan sebelumnya.

[4] melakukan penggerombolan daerah tertinggal dengan menggunakan metode Fuzzy K-Rataan, sehingga diperoleh 25 daerah agak tertinggal, 64 daerah h tertinggal, 79 daerah sangat tertinggal, dan 58 daerah sangat parah tertinggal. Gerombol daerah tertinggal dan sangat tertinggal tersebar merata di seluruh Indonesia, kecuali daerah di pulau Irian. Sebagian besar daerah di pulau Irian termasuk daerah yang agak tertinggal dan sangat parah tertinggal.

[5] mengelompokkan jumlah penumpang bus Trans Jogja berdasarkan jalur bus dan shelter dengan menggunakan metode clustering k-means, dan *Agglomerative Hierarchical Clustering (AHC)*. Penelitian ini menghasilkan 3 cluster untuk metode *K-Means*. Hasil pengelompokan metode *K-Means* memiliki hasil yang lebih baik dari pada metode AHC.

Tujuan dari penelitian ini adalah untuk menganalisis pengelompokan mahasiswa berdasarkan IPK dan IP pada kelompok MK (Mata Kuliah), dengan menggunakan analisis cluster tak berhirarki *K-means*. Selain itu juga bertujuan untuk menginterpretasikan pengaruh IP pada KBM (Kelompok Bidang Minat) terhadap pengelompokan dan penyebaran mahasiswa yang terbentuk, sehingga dapat dilihat kecenderungan (kemiripan/ketakmiripan) karakter IP mahasiswa pada suatu bidang ilmu (bidang minat).

II. KAJIAN LITERATUR

2.1 Metode Pengklasteran

Secara umum metode pengelompokan dalam analisis klaster dibedakan menjadi metode hirarki (*Hierarchical Clustering Method*) dan metode non hirarki (*Nonhierarchical Clustering Method*).

Dalam [3] metode hirarki dibedakan menjadi dua metode pengelompokan, yaitu:

1. Metode Aglomeratif
2. Metode Difusif

Pada metode non hirarki, Berdasarkan Hartigan (1975), data dibagi dalam k partisi, setiap partisi mewakili sebuah klaster.

2.2 Metode K-Means

Menurut [6], dasar pengelompokan metode *K-Means* adalah menempatkan objek berdasarkan rata-rata (*mean*) klaster terdekat. Oleh karena itu, metode ini bertujuan untuk meminimumkan *error* akibat

partisi objek ke dalam klaster. *Error* partisi disebut juga sebagai fungsi objektif.

2.3. Komponen K- Means

Algoritma *K- Means* memerlukan 3 komponen yaitu:

1. Jumlah Klaster k

2. Klaster Awal

Terdapat beberapa pendapat dalam memilih klaster awal untuk metode *K-Means* sebagai berikut:

- a. Berdasarkan [3], pemilihan klaster awal dapat ditentukan berdasarkan interval dari jumlah setiap observasi.
- b. Berdasarkan [7], pemilihan klaster awal dapat ditentukan melalui pendekatan salah satu metode hirarki.
- c. Berdasarkan [8], pemilihan klaster awal dapat secara acak dari semua observasi.

3. Ukuran Jarak

Dalam hal ini, ukuran jarak digunakan untuk menempatkan observasi ke dalam klaster berdasarkan centroid terdekat. Ukuran jarak yang digunakan dalam metode *K-Means* adalah jarak Euclid.

2.4 Algoritma K-Means

Menurut [3], *K-Means* adalah algoritma *clustering* yang berulang-ulang. Adapun algoritma *K-means* dalam pembentukan klaster sebagai berikut:

1. Misalkan diberikan matriks data $X = \{x_{ij}\}$ berukuran $n \times p$ dengan $i=1,2, \dots, n$, $j=1,2, \dots, p$ dan asumsikan jumlah klaster awal k .
2. Tentukan centroid.
3. Hitung jarak setiap objek ke setiap centroid dengan menggunakan jarak Euclid.
4. Setiap objek disusun ke centroid terdekat dan kumpulan objek tersebut akan membentuk klaster.
5. Tentukan centroid baru dari klaster yang baru terbentuk, di mana centroid baru itu diperoleh dari rata-rata setiap objek yang terletak pada klaster yang sama.
6. Ulangi langkah 3, jika masih ada objek dalam suatu klaster yang berpindah.

Perbedaan utama teknik *partitioning* dari teknik berhirarki adalah bahwa penempatan objek atau elemen ke dalam klaster-klaster mungkin berubah selama aplikasi algoritma, sementara pada teknik berhirarki sekali sebuah klaster telah ditemukan dan elemen telah ditempatkan dalam sebuah klaster, maka penempatan elemen ini tidak berubah [9].

III. METODE PENELITIAN

Penelitian ini bersifat studi kasus, dengan data yang digunakan pada penelitian ini yaitu data transkrip nilai mahasiswa Jurusan Matematika FMIPA Unsri angkatan 2012. Adapun tahapan-tahapan penelitian tersebut meliputi:

1. Pengumpulan data transkrip nilai.
2. Mendefinisikan MK dengan 3 kode yakni UNI untuk Mata Kuliah Umum (MKU), MIP untuk Mata Kuliah Dasar (MKD), serta MMP untuk Mata Kuliah Keilmuan Matematika (MKM) dan Kelompok Bidang Minat (KBM).
 1. Kelompok Bidang Minat (KBM) yaitu: Matematika Murni dan Komputasi (digabung), Optimasi, Statistika, dan Aktuaria.
 2. Kelompok MK PE, PIL, KKL, KKN digabung dan diberi nama Kelompok Mata Kuliah Pilihan (MKP).
3. Menyusun tabulasi nilai IP setiap mahasiswa berdasarkan MKU, MKD, MKM, MKP, KBM Matematika Murni dan Komputasi atau dinamakan KBM Murni, KBM Optimasi, KBM Statistika, KBM Aktuaria, jumlah SKS yang diambil pada setiap KBM, serta IPK, sehingga terbentuklah 13 variabel.
4. Entri data. Proses entri data dilakukan sesuai dengan nilai-nilai variabel yang telah didefinisikan.
5. Membuat matriks data X berukuran $n \times p$; dengan n adalah jumlah mahasiswa angkatan 2012 dan p adalah jumlah variabel yang diteliti.
6. Pengolahan data menggunakan analisis kluster digunakan untuk menganalisis kemiripan antar karakter mahasiswa ditinjau dari variabel-variabel yang telah diteliti dengan pendekatan non-hirarki, yaitu metode *K-Means*.
7. Standarisasi Matriks Data.
8. Melakukan analisis Komponen Utama untuk mencari skor komponen utama pertama dan komponen utama kedua.
9. Melakukan Analisis Biplot untuk mengetahui jumlah k kluster yang akan diambil.
10. Selanjutnya melakukan analisis kluster menggunakan metode *K-Means* dengan bantuan *software* IBM SPSS 23 dengan menggunakan jumlah kluster yang didapat dari langkah 9 yaitu dengan $k=2$, $k=3$, $k=4$, dan $k=5$.
11. Khusus untuk $k=3$ juga dikerjakan secara manual dengan langkah-langkah sebagai berikut:
 - a. Tentukan centroid
 - b. Hitung jarak setiap objek ke setiap centroid dengan menggunakan jarak Euclid.
 - c. Setiap objek disusun ke centroid terdekat dan kumpulan objek tersebut akan membentuk kluster.
 - d. Tentukan centroid baru dari kluster yang baru terbentuk, di mana centroid baru itu diperoleh dari

- rata-rata setiap objek yang terletak pada kluster yang sama.
- e. Ulangi langkah 3, jika masih ada objek atau anggota dalam suatu kluster yang berpindah.
12. Interpretasikan hasil.

IV. HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini adalah data mahasiswa Matematika Angkatan 2012

- Pada penelitian ini diasumsikan:
1. Kelompok MK wajib berkode UNI diberi nama Mata Kuliah Umum (MKU).
 2. Kelompok MK wajib berkode MIP diberi nama Mata Kuliah Dasar (MKD).
 3. Kelompok MK wajib berkode MMP diberi nama Mata Kuliah Keilmuan Matematika (MKM).
 4. Untuk MK pilihan yang berkode MMP terbagi menjadi 5 Kelompok Bidang Minat (KBM) yaitu: Matematika Murni, Komputasi, Statistika, Aktuaria, dan Optimasi.
 5. Untuk KBM Matematika Murni digabung dengan KBM Komputasi dan diberi nama KBM Murni.
 6. Untuk mata kuliah Pengantar Energi (PE), Pengantar Ilmu Lingkungan (PIL), Kuliah Kerja Lapangan (KKL), dan Kuliah Kerja Nyata (KKN) digabung menjadi satu kelompok mata kuliah yang diberi nama Mata Kuliah Pilihan (MKP).
 7. Peminatan atau kedominanan mahasiswa pada suatu KBM dilihat berdasarkan jumlah SKS yang diambilnya pada setiap KBM.
 8. Nilai IP dan IPK dikategorikan menjadi tinggi (T) dan rendah (R). Tinggi apabila IP atau IPK lebih besar dari 3 dan rendah apabila nilai IP atau IPK lebih kecil atau sama dengan 3.
 9. Tingkat kesesuaian pada tabel anggota kluster dilihat berdasarkan persentase berikut:

$T > 90\%$: Sangat sesuai
$75\% < T \leq 90\%$: Sesuai
$50\% < T \leq 75\%$: Tidak sesuai
$T \leq 50\%$: Sangat tidak sesuai
 10. Total mahasiswa Jurusan Matematika FMIPA Unsri angkatan 2012 adalah 65 mahasiswa. Total mahasiswa yang telah lulus pada Desember 2017 ada 57 mahasiswa.

Tabel 1 berikut ini merupakan statistik deskriptif dari variabel yang digunakan.

TABEL 1. STATISTIK DESKRIPTIF

	Minimum	Maximum	Mean	Std. Deviation
MKU	2,67	4,00	3,4326	0,29272
MKD	2,30	3,75	2,9868	0,29528
MKM	2,45	3,77	2,8823	0,26403
MKP	2,50	4,00	3,4621	0,40863
JSM	0	15	5,74	4,402
Murni	0,00	4,00	2,7072	1,40101
JSO	3	24	13,32	4,318

Optimasi	2,33	4,00	3,1718	0,46977
JSS	0	18	8,95	4,518
Statistika	0,00	4,00	2,8523	0,78416
JSA	0	12	4,42	2,542
Aktuaria	0,00	4,00	3,2149	0,88064
IPK	2,64	3,80	3,0240	0,23435

Data yang digunakan memiliki jangkauan yang berbeda, oleh sebab itu dilakukan standarisasi data dengan mengubah data menjadi *Z-score*.

Sebelum melakukan proses analisis kluster, terlebih dahulu dilakukan pendeteksian terhadap *outlier* dan multikolonieritas.

Uji *outlier* dilakukan dengan menentukan nilai batas yang akan dikategorikan sebagai data *outlier* yaitu dengan cara mengkonversi nilai data kedalam *Z-score*, yang memiliki nilai rata-rata (*means*) sama dengan nol dan standar deviasi sama dengan satu.

Oleh karena data *outlier* ini merupakan representasi dari populasi yang diteliti maka *outlier* tersebut tetap dipertahankan. Jika 4 observasi dikeluarkan, berpengaruh terhadap matriks korelasi.

Tabel 2 berikut ini merupakan nilai skor outlier.

TABEL2. NILAI SKOR OUTLIER

Observasi	Z _{MKD}	Z _{MKM}	Z _{IPK}	Observasi	Z _{JSA}
Iffah	2,5845	3,3621	3,3111	Novaldiyu	2,9813
Alvian	2,5845	3,0970	2,7564	Kikiwiyandi	2,9813

Oleh karena data *outlier* ini merupakan representasi dari populasi yang diteliti maka *outlier* tersebut tetap dipertahankan. Jika 4 observasi dikeluarkan, berpengaruh terhadap matriks korelasi.

Cara mengatasi multikolonieritas adalah dengan analisis komponen utama. Namun, karena disini kita memang menginginkan adanya hubungan antara kluster yang terbentuk dengan IP dan IPK maka multikolonieritas diabaikan. Jika dilihat nilai korelasi pada Tabel 3, IPK dengan peubah yang lain, maka IPK paling berkorelasi kuat dengan peubah MKM.

AKU digunakan untuk mereduksi data asal menjadi komponen utama (KU) yang jumlahnya sedikit mungkin tanpa banyak kehilangan informasi data asal.

TABEL 3. ENTRI DARI MATRIKS KORELASI

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
X1	1	0,37	0,34	0,15	0,37	0,22	0,02	0,30	0,29	0,28	0,03	0,03
X2	0,37	1	0,53	0,06	0,04	0,02	0,09	0,28	0,09	0,25	0,04	0,05
X3	0,34	0,53	1	0,10	0,36	0,18	0,09	0,57	0,21	0,51	0,20	0,29
X4	0,15	0,06	0,10	1	0,20	0,01	0,11	0,07	0,01	0,19	0,19	0,05
X5	-0,37	0,04	0,36	0,20	1	0,55	0,16	0,43	0,70	0,38	0,34	0,16
X6	-0,22	0,02	0,18	0,01	0,55	1	0,25	0,18	0,47	0,14	0,43	0,15
X7	0,02	0,09	0,09	0,11	0,16	0,25	1	0,12	0,47	0,21	0,39	0,16
X8	0,30	0,28	0,57	0,07	0,43	0,18	0,12	1	0,28	0,36	0,11	0,28
X9	0,29	0,09	0,21	0,01	0,70	0,47	0,47	0,28	1	0,39	0,13	0,13
X10	0,28	0,25	0,51	0,19	0,38	0,14	0,21	0,36	0,39	1	0,27	0,21
X11	-0,03	0,04	0,20	0,19	0,34	0,43	0,39	0,11	0,13	0,27	1	0,24
X12	-0,03	0,05	0,29	0,05	0,16	0,15	0,16	0,28	0,13	0,21	0,24	1
X13	0,49	0,64	0,95	0,19	0,39	0,17	0,07	0,69	0,21	0,55	0,17	0,26

Tahap awal menentukan KU adalah mendapatkan nilai eigen dan vektor eigen dari matriks korelasi **R**, yang entri-entrinya seperti pada Tabel 3.

Analisis Komponen Utama (AKU) menghasilkan kombinasi linear dua Komponen Utama (KU) dengan masing-masing peubah, yaitu sebagai berikut:

$$KU1 = 0,13x_1 + 0,11x_2 + 0,19x_3 + 0,05x_4 - 0,16x_5 - 0,11x_6 - 0,05x_7 + 0,16x_8 + 0,14x_9 + 0,16x_{10} + 0,09x_{11} + 0,08x_{12} + 0,20x_{13}$$

$$KU2 = 0,10x_1 + 0,27x_2 + 0,19x_3 - 0,01x_4 + 0,18x_5 + 0,25x_6 + 0,30x_7 + 0,07x_8 - 0,23x_9 - 0,23x_{11} - 0,07x_{12} + 0,21x_{13}$$

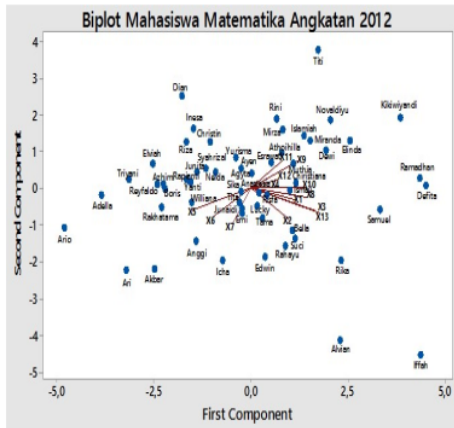
Berdasarkan dua komponen utama pertama dapat digambarkan variabel sebagai vektor dan juga skor komponen dari objek sebagai titik-titik pada biplot pada Gambar 1.

Ukuran kesesuaian dari biplot yang diperoleh sebagai berikut:

$$\rho_2 = \frac{\lambda_1 + \lambda_2}{\sum_{k=1}^5 \lambda_k} = \frac{42962166}{429621661123090007970631059504940347026701140005} = \frac{6,462}{12,999}$$

$$\rho_2 = 0,50$$

Nilai ρ_2 yang dihasilkan adalah 0,50 maka dapat diinterpretasikan bahwa biplot ini dapat mewakili keragaman data hanya sebesar 50%.



Gambar 1. Biplot mahasiswa Matematika angkatan 2012 dan karakteristiknya

Berdasarkan Gambar 1, mahasiswa angkatan 2012 dapat dikelompokkan menjadi banyak kluster. Oleh karena itu untuk keperluan analisis selanjutnya maka data pada lampiran 2 akan dikelompokkan menjadi 2 kelompok, 3 kelompok, 4 kelompok, dan 5 kelompok.

Untuk pemilihan nilai $k=2$, $k=3$, $k=4$, dan $k=5$ menggunakan bantuan software IBM SPSS 23, khusus untuk $k=3$ dikerjakan juga dengan cara manual. Berikut ini adalah hasil *output* dari software IBM SPSS 23 menggunakan masing-masing nilai k tersebut.

Berdasarkan kluster $k=2$, $k=3$, $k=4$, dan $k=5$ dapat dilihat bahwa kluster dengan $k=3$ mempunyai karakteristik tiap kluster yang heterogen dan memiliki karakteristik yang homogen pada tiap anggota dalam kluster serta anggota-anggota tiap kluster memiliki kesesuaian dengan karakteristik kluster tersebut sebesar 100% atau dengan kata lain kluster yang terbentuk sangat sesuai dengan data. Oleh karena itu, untuk keperluan analisis kluster *K-Means* selanjutnya secara manual, diambil $k=3$.

Data dikelompokkan menjadi 3 kluster, $k = 3$, dipilih data ke-1, ke-40, dan ke-27 sebagai pusat kluster awal, dapat dilihat pada Tabel 4.

Tabel 4 merupakan tampilan pertama kluster data sebelum dilakukan iterasi. Langkah selanjutnya menentukan pusat kluster baru dengan menghitung jarak *euclid* antara data ke-1 dengan pusat awal kluster 1, 2, dan 3. Dari hasil perhitungan jarak antara setiap data terhadap semua pusat kluster, didapatkan nilai jarak terkecil terhadap satu pusat kluster, maka pusat kluster tersebut disebut sebagai pusat kluster

terdekat, dan data akan berafiliasi menjadi kluster dari pusat kluster terdekat.

TABEL 4. PUSAT KLUSTER AWAL

	Kluster		
	1	2	3
X ₁	1,35751	0,-89722	1,35751
X ₂	2,58453	-1,81808	1,39921
X ₃	3,36213	-0,08439	-1,63721
X ₄	0,09274	0,50876	0,09274
X ₅	-0,62175	-1,30329	1,42286
X ₆	0,92277	-1,93231	0,38744
X ₇	1,77942	-1,69410	1,08471
X ₈	1,76308	0,69873	-1,42996
X ₉	,01165	1,33982	-1,31651
X ₁₀	1,46364	0,69848	-1,08688
X ₁₁	-0,55900	1,80123	-1,73912
X ₁₂	-0,24404	0,89149	-3,65064
X ₁₃	3,31111	-0,05989	-0,99865

Jarak antara anggota kluster dapat dihitung berdasarkan entri Tabel 4.

$$d(x_1, c_1) = \sqrt{(1,35751-1,35751)^2 + \dots + (3,31111-3,31111)^2} = 0$$

$$d(x_1, c_2) = \sqrt{(1,35751-1,35751)^2 + \dots + (3,31111-(-0,99865))^2} = 8,88285$$

$$d(x_1, c_3) = \sqrt{(1,35751-(-0,89722))^2 + \dots + (3,31111-(-0,05989))^2} = 9,02133$$

Berdasarkan perhitungan jarak di atas dapat diketahui bahwa data ke-1 lebih dekat jaraknya terhadap pusat kluster ke-1 sehingga dapat disimpulkan data ke-1 masuk ke kluster 1. Untuk perhitungan 56 data berikutnya menggunakan bantuan program Microsoft Excel 2010. Setelah dipastikan data 1 sampai 57 sudah masuk ke salah satu kluster. Selanjutnya tentukan pusat kluster baru berdasarkan data yang tergabung pada setiap klasternya. Pengklasteran dihentikan apabila tidak ada lagi data yang berpindah, sehingga didapatkan pusat kluster akhir pada Tabel 5 berikut.

TABEL 5. PUSAT KLUSTER AKHIR

	Kluster		
	1	2	3
X ₁ (MKU)	0,85191	-0,09212	-0,33856
X ₂ (MKD)	1,00975	-0,15864	-0,31402
X ₃ (MKM)	1,42678	-0,05661	-0,73938
X ₄ (MKP)	0,33746	-0,04267	-0,12319
X ₅ (JSM)	-0,34914	-0,32642	0,78141
X ₆ (Mumi)	-0,02869	-0,18167	0,33748
X ₇ (JSO)	0,66789	-0,58257	0,63520
X ₈ (Optimasi)	1,07338	0,16372	-0,92032
X ₉ (JSS)	-0,12116	0,43224	-0,69149
X ₁₀ (Statistika)	0,97011	0,16883	-0,86858
X ₁₁ (JSA)	0,03106	0,38509	-0,69784
X ₁₂ (Aktuaria)	0,43728	0,29534	-0,77841
X ₁₃ (IPK)	1,55307	-0,05704	-0,81290

fa

Berdasarkan Tabel 5 dapat didefinisikan sebagai berikut:

1. Klaster 1

Dalam klaster 1 berisikan mahasiswa-mahasiswa yang mempunyai IP kelompok Mata Kuliah Umum (MKU) tinggi, IP kelompok Mata Kuliah Dasar (MKD) tinggi, IP kelompok Mata Kuliah Keilmuan Matematika (MKM) tinggi, IP kelompok Mata Kuliah Pilihan (MKP) tinggi, IP KBM Optimasi tinggi, IP KBM Statistika tinggi, serta IP KBM Aktuaria tinggi dan merupakan mahasiswa yang lebih dominan mengambil mata kuliah dari KBM Optimasi atau dapat dikatakan bahwa semua kelompok MK tersebut nilai IP nya lebih dari rata-rata nilai IP populasi mahasiswa yang diteliti. Hal ini terbukti dari nilai positif (+) yang terdapat pada tabel pusat klaster akhir dalam keseluruhan variabel. Dengan demikian dapat diduga bahwa klaster 1 merupakan pengelompokan dari mahasiswa-mahasiswa yang mempunyai IPK tinggi dengan rata-rata IPK yaitu 3,39.

2. Klaster 2

Karakteristik mahasiswa yang masuk dalam pengelompokan klaster 2 yaitu mahasiswa yang memiliki IP KBM Optimasi, IP KBM Statistika dan IP KBM Aktuaria melebihi rata-rata dari populasi mahasiswa yang diteliti, serta merupakan kelompok mahasiswa yang lebih dominan mengambil MK dari KBM Statistika dan Aktuaria. Untuk kelompok MK yang lain mahasiswa-mahasiswa di klaster 2 ini nilai IP nya berada di bawah rata-rata nilai IP populasi mahasiswa yang diteliti. Dengan demikian dapat diduga bahwa mahasiswa dengan IPK sedang berada pada klaster 2 dengan rata-rata IPK yaitu 3,01.

3. Klaster 3

Karakteristik mahasiswa-mahasiswa yang masuk dalam pengelompokan klaster ini adalah mahasiswa yang hampir seluruh IP kelompok MK atau hampir seluruh variabel kelompok MK nilai IP nya berada di bawah rata-rata nilai IP populasi mahasiswa yang diteliti, kecuali IP KBM Murni paling tinggi diantara klaster yang lain dan merupakan kelompok mahasiswa yang lebih dominan mengambil mata kuliah dari KBM Murni. Dengan demikian dapat diduga bahwa klaster 3 merupakan kelompok mahasiswa yang memiliki IPK rendah dengan rata-rata IPK yaitu 2,83.

Anggota tiap-tiap klaster adalah sebagai berikut:

1. Klaster 1

Iffah Husniah, Ramadhan Pratama, Defita Yolanda, Rahayu Tami Agustin, Samuel Hutapeah, M. Edwin Afriansyah, Suci Novrita

KD, Bella Arisha, Rika Apriani, dan Alvian Yayang Hilman.

2. Klaster 2

Emi Widarti, Christiana BR Hotang, Agyta Meitriova, Atoihilla Abdul Latif, Elinda Damayanti, Christin Nova Samosir, Islamiah, Esrawati Pardosi, Mirza Denia Putri, Ayen Ruth Borus, Dewi Rakhmatia Nur, Dian Permata Sari, Lucky Sri Dorce, Sika Amelia, Risfa Risa Octa R, Tita Adestiy, Rini Ade Sitinjak, Nelda Amelia, Titi Larastiana, Ismail, Anastasia Carolina, M.Albar Pratama, Muthia Ulfa, Ahmad Junaidi, Inesa Larasati, Riza Rahma Pratiwi, Miranda Avivana, Novaldiyu Pratama, Yurisma Nandasari, dan Kiki Wiyandi.

3. Klaster 3

Icha Puspita Novyasti, Anggi Nurul Pratiwi, Azhimi, Muhammad Syahrizal, Rakhata Gusril, Junita Sihombing, M. Ari Juliansyah, Yanti Wulandari, Elvia Anggraini, Reyfaldo Tomi, Rapianti, Adella Rosita, Wiliana Gita Putri, Boris Tambun, Muhammad Ario W, Akbar Yulanda, dan Triyani.

IV. KESIMPULAN

Berdasarkan hasil dan pembahasan, maka dapat ditarik kesimpulan sebagai berikut:

1. Dengan menggunakan analisis biplot didapat 4 nilai k yaitu $k=2$, $k=3$, $k=4$, dan $k=5$. Klaster $k=3$ sangat sesuai dengan data, klaster $k=2$ dan $k=4$ sesuai dengan data, sedangkan klaster $k=5$ tidak sesuai dengan data, sehingga dipilih klaster $k=3$ karena sangat sesuai dengan data. Jadi didapatkan 3 klaster atau 3 kelompok mahasiswa berdasarkan nilai IP pada kelompok MK dan jumlah SKS yang diambil pada setiap KBM, serta IPK. IPK yang tinggi direpresentasikan oleh IP kelompok mata kuliah yang juga tinggi.
2. Karakteristik ketiga klaster adalah:

a) Klaster 1

Mahasiswa yang termasuk ke dalam klaster 1 berjumlah 10 mahasiswa. Klaster 1 mempunyai karakteristik nilai IP kelompok Mata Kuliah Umum (MKU), Mata Kuliah Dasar (MKD), Mata Kuliah Keilmuan Matematika (MKM), Mata Kuliah Pilihan (MKP), KBM Optimasi, KBM Statistika, serta KBM Aktuaria, nilai IP kelompok mata kuliahnya melebihi rata-rata nilai IP populasi mahasiswa yang diteliti dan merupakan mahasiswa yang lebih dominan mengambil mata kuliah dari KBM Optimasi serta merupakan kelompok mahasiswa yang memiliki IPK tinggi dengan rata-rata IPK 3,39.

b) Klaster 2

Mahasiswa yang termasuk ke dalam klaster 2 berjumlah 30 mahasiswa. Karakteristik

mahasiswa pada klaster 2 yaitu memiliki nilai IP KBM Optimasi, KBM Statistika dan KBM Aktuaria, nilai IP kelompok MK nya melebihi rata-rata dari IP populasi mahasiswa yang diteliti, sementara variabel kelompok MK yang lain berada di bawah rata-rata dan merupakan kelompok mahasiswa yang lebih dominan mengambil MK dari KBM Statistika dan Aktuaria serta merupakan kelompok mahasiswa yang mempunyai IPK sedang dengan rata-rata IPK yaitu 3,01.

c) Klaster 3

Mahasiswa yang termasuk ke dalam klaster 3 berjumlah 17 mahasiswa. Klaster 3 merupakan mahasiswa yang hampir seluruh nilai IP kelompok MK-nya berada di bawah rata-rata IP populasi mahasiswa yang diteliti kecuali nilai IP KBM Murni paling tinggi diantara klaster yang lain dan merupakan kelompok mahasiswa yang mempunyai IPK rendah dengan rata-rata IPK 2,83 serta KBM yang dominan diambil yaitu KBM murni.

REFERENSI

- [1] M. Saleh, "Pengaruh Motivasi, Faktor Keluarga, Lingkungan Kampus, dan Aktif Berorganisasi Terhadap Prestasi Akademik," *J. Phenom.*, vol. 4, no. 2, pp. 100–111, 2014.
- [2] Irmeilyana, R. Sania, A. Desiani, H. Tanuji, "Peng-cluster-an Mahasiswa Berdasarkan IPK dan IP Kelompok Mata Kuliah dan Kelompok Bidang Minat," in *ARS 2017*, 2017.
- [3] J. A. Hartigan, *Clustering Algorithms*. New York: John Wiley & Sons, Inc., 1975.
- [4] T. Agustin and A. Djuraidah, "Penggerombolan daerah tertinggal di Indonesia dengan fuzzy k-rataan," *Forum Stat. dan Komputasi*, vol. 15, no. 1, pp. 22–27, 2010.
- [5] L. Zahrotun, "Analisis pengelompokan jumlah penumpang bus Trans Jogja menggunakan metode clustering k-means dan agglomerative hierarchical clustering (AHC)," *J. Inform.*, vol. 9, no. 1, pp. 1039–1047, 2015.
- [6] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, 6th ed. New Jersey: Prentice-Hall, Inc., 2007.
- [7] A. C. Rencher, *Methods of Multivariate Analysis*. Toronto: John Wiley & Sons, Inc., 2002.
- [8] S. Sharma, *Applied Multivariate Techniques*. New York: John Wiley & Sons, Inc., 1996.
- [9] Härdle & Simar, *Applied Multivariate Statistical Analysis*, 2nd ed. New York: Springer Berlin Heidelberg, 2007.

2018-ARS-K-means-Kelly

ORIGINALITY REPORT

10%

SIMILARITY INDEX

9%

INTERNET SOURCES

5%

PUBLICATIONS

5%

STUDENT PAPERS

PRIMARY SOURCES

1	repository.radenfatah.ac.id Internet Source	1%
2	kc.umn.ac.id Internet Source	1%
3	repository.uinjambi.ac.id Internet Source	1%
4	siepub.unsri.dev Internet Source	1%
5	fmipa.unmul.ac.id Internet Source	1%
6	jurnal.poltekba.ac.id Internet Source	1%
7	Submitted to Syiah Kuala University Student Paper	1%
8	aan-themax.blogspot.com Internet Source	1%
9	Agnes Mariana Veronica Sandag, Charles E. Mongi, Mans L. Mananohas. "Pengelompokan Sekolah Menengah Atas (SMA) di Kabupaten	1%

Minahasa Tenggara Berdasarkan Standar Kompetensi Lulusan Tahun 2018 Menggunakan Analisis Gerombol", d'CARTESIAN, 2021

Publication

10	Submitted to Universitas Diponegoro Student Paper	1 %
11	doaj.org Internet Source	1 %
12	journal.universitاسbumigora.ac.id Internet Source	1 %

Exclude quotes Off

Exclude matches < 1%

Exclude bibliography Off