

Analisis Cluster Mahasiswa Pendidikan Matematika Universitas Nias

Memori Putra Lawuna^{1✉}, Yulisman Zega², Ratna Natalia Mendrofa³

^{1,2,3} Program Studi Pendidikan Matematika, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Nias,
Jl. Yos Sudarso Ujung No.118/E-S, Ombolata Ulu, Kec. Gunungsitoli, Kota Gunungsitoli, Sumatera Utara
memorylawuna@gmail.com

Abstract

Data is true and real information and can be used as a basis for study or information. To get accurate information, the data must be processed first. This research aims to process student data from the Mathematics Education Study Program at Nias University using one of the clustering methods, namely the K-Means Cluster Method. The data used in this research is primary data obtained directly from respondents using a questionnaire and then processed using IBM SPSS Statistics 26. The processing results in applying the cluster analysis method produce 4 clusters with the number of each cluster member being cluster 1 with 45 student data, cluster 2 has 29 student data, cluster 3 has 65 student data, and cluster 4 has 64 student data. Semester 8 students all fall into the same cluster compared to students in other semesters. This is due to the influence of the number of SKS taken. Students with a high IPK enter Cluster 2 with an average of 3.37. Students with a moderate IPK enter Cluster 4 with an average of 3.27, students with a low IPK enter Cluster 1 with an average of 2.98, and students with a very low IPK enter Cluster 3 with an average of 2.93. The clusters formed are said to be good where the clusters have significant differences as evidenced by a significant ANOVA test value of < 0.05 and the variables that have the most influence in the formation of clusters, respectively, seen from the F value, are number of SKS, IPS, semester, scholarship, IPK, age, district, place living, owning a laptop, owning a vehicle, parent's income, average duration of study, and finally school origin.

Keywords: Student Data, Cluster Analysis, K-Means

Abstrak

Data adalah informasi yang akurat dan asli yang dapat menjadi landasan penelitian atau informasi lainnya. Untuk mendapatkan informasi yang akurat, Pertama, data harus diproses. Tujuan dari penelitian ini adalah untuk mengolah data mahasiswa Program Studi Pendidikan Matematika Universitas Nias dengan salah satu metode *clustering* yakni *K-Means Cluster Method*. Data primer dari responden dikumpulkan secara langsung dengan menggunakan kuesioner dan digunakan dalam penelitian ini dan kemudian diolah dengan IBM SPSS *Statistic* 26. Hasil pengolahan dalam penerapan metode *cluster analysis* menghasilkan 4 *cluster* dengan jumlah tiap anggota *cluster* ialah Data mahasiswa tersedia 45 mahasiswa di cluster 1, 29 mahasiswa di cluster 2, 65 mahasiswa di cluster 3, dan 64 mahasiswa di cluster 4. Mahasiswa semester 8 Jika dibandingkan dengan mahasiswa semester sebelumnya, mereka semua berada dalam satu cluster yang sama. Hal ini disebabkan pengaruh jumlah SKS yang diambil. Mahasiswa dengan IPK yang tinggi masuk ke Cluster 2 dengan rata-rata 3.37. Mahasiswa dengan IPK yang sedang masuk ke Cluster 4 dengan rata-rata 3.27, Mahasiswa dengan IPK yang rendah masuk ke Cluster 1 dengan rata-rata 2.98, dan Mahasiswa dengan IPK Yang terendah termasuk dalam Cluster 3 dengan rata-rata 2,93. *Cluster* yang terbentuk dikatakan baik dimana cluster memiliki perbedaan signifikan dibuktikan dengan nilai uji ANOVA signifikan < 0.05 dan faktor-faktor yang, pada gilirannya, mempunyai dampak terbesar terhadap pembentukan klaster dilihat dari nilai F adalah Jumlah SKS, IPS, Semester, Beasiswa, IPK, Usia, Kabupaten, Tempat Tinggal, Memiliki Laptop, Memiliki Kendaraan, Pendapatan Orangtua, Rata-rata durasi belajar dan yang terakhir Asal Sekolah.

Kata kunci: Data Mahasiswa, Analisis Cluster, K-Means

Copyright (c) 2024 Memori Putra Lawuna, Yulisman Zega, Ratna Natalia Mendrofa

✉ Corresponding author: Memori Putra Lawuna

Email Address: memorylawuna@gmail.com (Jl. Yos Sudarso Ujung Ombolata Ulu, Sumatera Utara)

Received 04 November 2023, Accepted 12 February 2024, Published 12 February 2024

DoI: <https://doi.org/10.31004/cendekia.v8i2.2940>

PENDAHULUAN

Pulau Nias saat ini memiliki beberapa Perguruan Tinggi dimana salah satunya adalah Universitas Nias. Universitas Nias adalah satu dari beberapa perguruan tinggi yang berada di Pulau Nias yang terbentuk dari 3 Fakultas yakni Fakultas Keguruan dan Ilmu Pendidikan, Fakultas

Ekonomi dan Fakultas Saintek. “Universitas Nias didirikan pada tanggal 22 September 2021 berdasarkan SK Kemdikbud Ristekdikti Nomor 400/E/O/2021. Universitas Nias merupakan penggabungan Perguruan Tinggi IKIP Gunungsitoli menjadi Fakultas Keguruan dan Ilmu Pendidikan, STIE Pembnas menjadi Fakultas Ekonomi serta penambahan fakultas baru yakni Fakultas Sains dan Teknologi. Pada Fakultas Keguruan dan Ilmu Pendidikan terdapat salah satu program studi yakni program studi pendidikan matematika dimana yang saat ini juga calon peneliti berada”. Mahyudi mengatakan dalam Zega, (2022), “Matematika merupakan bidang studi yang memiliki peranan penting dalam bidang pendidikan yang mendunia”. Dalam Program Studi ini ada banyak data-data yang dapat diolah untuk dijadikan referensi kedepan dalam meningkatkan program studi. Salah satu contohnya adalah data mahasiswa.

Mengolah data yang besar dengan jumlah record dan atribut yang banyak bukanlah hal yang gampang. Data mining adalah salah satu teknik yang dapat digunakan untuk memproses data dalam jumlah besar. Teknik data mining ialah suatu alat agar dapat mengekstrak data dari database yang besar dengan spesifikasi yang kompleks. Karena Informasi yang komprehensif adalah salah satu tuntutan agar data mining itu sendiri dapat dipenuhi dengan kemungkinan dan prosedur, dan informasi yang diperoleh dapat dipenuhi digunakan untuk pengambilan keputusan atau untuk menentukan kualitas pengambilan keputusan. Pengumpulan data atau informasi dapat dijadikan kesimpulan untuk mengambil keputusan dengan menggunakan cara menganalisis dan menelaah informasi yang terkandung dalam data tersebut (Lestari, 2019).

Kekayaan data mahasiswa dapat diolah menjadi data yang berguna bagi universitas untuk mengetahui informasi yang tersembunyi. Pengolahan data mahasiswa diperlukan untuk mengungkap informasi yang relevan berupa pengetahuan baru (knowledge Discovery), seperti rincian cara mengorganisasikan data siswa sesuai dengan tempat asal mahasiswa. Universitas dapat menggunakan informasi baru ini untuk meningkatkan kualitas mahasiswanya. Penerapan teknik data mining pada data siswa untuk menarik kesimpulan merupakan salah satu cara untuk meningkatkan kualitas mahasiswa. (Marajari et al., 2021).

Menurut (Nasari, 2015) “Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan”. Dimana sesuai (Maukar et al., 2022) “Knowledge Discovery in Database (KDD) adalah tahapan global dari penggalian data yang terdiri dari beberapa langkah / tahapan, diantaranya pemilihan data mentah, pra pengolahan data (preprocessing), transformasi data, data mining, dan evaluasi hasil dari pengolahan informasi”. Sedangkan menurut (Parlambang, 2020) “Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (machine learning) untuk menganalisis dan mengekstraksi pengetahuan (knowledge) secara otomatis”.

Banyaknya mahasiswa prodi pendidikan matematika adalah lebih dari 200 mahasiswa dengan tingkatan semester yang berbeda. Data Mahasiswa Matematika Universitas Nias dalam penelitian ini adalah berisi data Kabupaten/kota Mahasiswa berasal, Jurusan yang diambil pada saat SMA/SMK, jenis kelamin, semester, usia, rata-rata waktu harian yang dihabiskan untuk belajar mandiri, jumlah Satuan Kredit Semester (SKS), nilai Indeks Prestasi Semester (IPS), dan nilai Indeks Prestasi Kumulatif (IPK) pada semester terakhir mahasiswa. Proses belajar yang tinggi, yang peneliti gambarkan sebagai lama studi, diperlukan untuk menunjang variabel IPK dan IPS, yang merupakan penilaian terhadap hasil belajar yang dicapai mahasiswa pada setiap akhir semester. Hal ini akan membantu siswa mendapatkan hasil maksimal dalam penilaian tersebut. Dalam hal berapa lama waktu belajar, sebagian besar siswa melakukan kesalahan. Banyak anak mengabaikan fakta bahwa berprestasi di sekolah juga membutuhkan banyak pekerjaan. Dengan demikian, pengelompokan siswa menurut rata-rata kelompok lama belajar dan usia mahasiswa menjadi hal yang menarik bagi peneliti. Selain itu, faktor lain - penerima beasiswa dan pendapatan orang tua - ditambahkan sebagai variabel dalam penelitian ini.

Tujuan pengolahan data mahasiswa program studi pendidikan matematika di Universitas Nias khususnya menggunakan data mahasiswa dan teknik K-Means Clustering. Data yang digunakan ialah data mahasiswa semester 2 sampai dengan semester 8 dan sumber data berasal dari kuesioner yang telah disebar secara online. Manfaat bagi khalayak umum salah satunya ialah untuk mengetahui persebaran mahasiswa yang ada di program studi pendidikan matematika dan juga sebagai referensi kedepannya bagi program studi dalam mengembangkan program studi matematika baik dari segi kuantitas maupun kualitas mahasiswanya.

Data yang terkumpul tidak dapat memberikan hasil terbaik jika bukan dianalisis dengan baik. Alhasilnya, kita dapat menggunakan salah satu strategi tersebut. metode data mining yang umum digunakan yaitu *clustering*. Analisis klaster saat ini mulai diterapkan secara luas di beberapa domain keilmuan. Menurut (Riana, 2021) “Analisis cluster merupakan suatu teknik yang digunakan untuk menghasilkan objek atau kasus (responden) ke dalam kelompok yang relative homogen, yang disebut cluster. Objek/kasus dalam setiap kelompok cenderung mirip satu sama lain dan berbeda jauh (tidak sama) dengan objek dari cluster lainnya”.

(Awaliah, 2018) menyatakan bahwa “Cluster-cluster yang baik adalah cluster yang terbentuk memiliki Homogenitas (kesamaan) internal yang tinggi antar anggota dalam satu cluster (Within-Cluster), dan Heterogenitas (perbedaan) eksternal yang tinggi antar anggota dalam satu cluster (Between-Cluster)”.

Dari pemaparan (Rokhimah & Kismiantini, 2022), “Analisis *cluster* merupakan teknik multivariat yang digunakan untuk mengklasifikasi objek atau kasus (responden) ke dalam kelompok yang relatif homogen, yang disebut *cluster*. Objek atau kasus dalam setiap kelompok cenderung mirip satu sama lain dan berbeda jauh atau tidak samadengan objek dari *cluster* lainnya”. Adapun menurut (Nisa, 2019) “Analisis cluster bermanfaat dalam menyimpulkan suatu data yang kompleks

dengan cara mengelompokkan objek-objek yang memiliki kemiripan karakteristik”. Berdasarkan pemaparan (S.Y., 2019) “Terdapat dua jenis data clustering yang sering dipergunakan dalam proses pengelompokan data yaitu Hierarchical dan Non-Hierarchical, dan K-Means merupakan salah satu metode data clustering non-hierarchical atau Partitional Clustering”.

Menurut (Ayuningtias, 2019) “Perbedaan antara analisis cluster hirarki dan analisis cluster non-hirarki terletak pada awal untuk memulai pengelompokkan. Pada analisis cluster hirarki dimulai dengan dikelompokkannya dua atau lebih objek yang kesamannya paling dekat sehingga cluster akan membentuk seperti pohon yang terdapat tingkatan jelas antar objek, seperti dari objek yang mempunyai kesamaan dekat hingga yang paling tidak dekat. Berbeda dengan metode hirarki, analisis cluster non-hirarki dimulai dengan terlebih dahulu menentukan jumlah cluster yang diinginkan peneliti”.

Uraian dari Lestari, (2018) “salah satu metode non-hierarki yaitu metode K-means yang digolongkan sebagai metode pengklasifikasian yang bersifat unsupervised (tanpa arahan) dikarenakan data yang di analisis tidak mempunyai label kelas yang berarti dalam proses pengelompokannya, analisis ini tidak mempunyai anggota cluster yang pasti. Melainkan data yang telah masuk ke dalam cluster bisa saja berpindah ke cluster yang lain. Akan tetapi juga, karena peneliti sering menentukan sendiri jumlah cluster awal, baik dengan menggunakan metode tertentu atau berdasarkan pengalaman, maka metode K-Means ini sering disebut juga sebagai metode semi-supervised classification”. Menurut Nugroho, (2022) “Penggunaan K-Means Cluster Analysis untuk mengorganisasikan data merupakan salah satu pendekatan analisis cluster non-hierarchical yaitu strategi cluster dimana jumlah cluster diatur secara manual. Tujuan dari cluster ini adalah untuk membagi item yang diamati menjadi satu atau lebih kelompok tergantung pada karakteristiknya”.

Pada penerapan metode K-Means Cluster Analysis menurut Puntoriza (2020) “data yang bisa diolah dalam perhitungan adalah data numerik yang berbentuk angka. Setiap data dihitung kedekatan dengan nilai centroid yang sudah ditentukan sebelumnya, jarak terkecil antara data dengan masing-masing centroid merupakan anggota cluster yang terdekat”. (Muliono, 2019) “K-Means adalah salah satu algoritma dari teknik data mining yang mampu melakukan klasterisasi terhadap data heterogen karena pada dasarnya algoritma pengelompokkan hanya mampu mengenali nilai atribut homogen saja”.

METODE

Penelitian ini menggunakan desain penelitian terapan dengan metodologi kuantitatif. Penggunaan metode ilmiah dalam penelitian terapan menjamin objektivitas dalam pengumpulan fakta dan menguji kreatifitas. Data primer dikumpulkan langsung dari responden melalui kuesioner dan disusun untuk digunakan dalam penelitian ini. Dengan memanfaatkan SPSS, hasil tabulasi data diolah. Variabel penelitiannya meliputi X_1 (Kabupaten/kota Mahasiswa berasal), X_2 (Asal Sekolah Mahasiswa), X_3 (Usia mahasiswa), X_4 (Semester mahasiswa), X_5 [Rata-rata waktu harian yang

dihabiskan siswa untuk belajar mandiri (Jam)], X_6 (Jumlah SKS mahasiswa pada semester terakhir), X_7 (IPS mahasiswa pada semester terakhir), X_8 (IPK mahasiswa pada semester terakhir), X_9 (Memiliki Laptop), X_{10} (Tempat tinggal mahasiswa sekarang), X_{11} [Memiliki Kendaraan (Sepeda motor)], X_{12} (Penerima Beasiswa), X_{13} (Pendapatan Orang Tua Mahasiswa).

Dalam penelitian ini yang menjadi Seluruh mahasiswa program studi pendidikan matematika Universitas Nias dijadikan sampel penelitian semester 2 sampai dengan semester 8 dengan total responden 203 mahasiswa. Pengumpulan data dilakukan menggunakan *Google Form* dengan menyebarkan kuesioner kepada siswa sebagai responden melalui grup *WhatsApp* atau menghubungi responden secara langsung. Perangkat lunak pengolahan data yang digunakan adalah IBM SPSS Statistics 26. Analisis data yang dipergunakan ialah 1) Mengumpulkan instrumen yang telah diterima dan diselesaikan oleh responden kemudian diperiksa kelengkapannya, 2) Gunakan *SPSS Data View* dan *Variable View* untuk mengatur semua variabel sebagai aturan utama sistem data numerik, 3) Masukkan informasi instrumen pada tabel data penelitian, 4) Melakukan uji asumsi multikolinearitas terhadap data, 5) Melakukan standarisasi data, 6) Tetapkan protokol untuk analisis kluster. Penelitian ini menggunakan K-Means Cluster, sebuah teknik analisis cluster non-hierarki, 7) Jalankan hasil analisis cluster, 8) Setelah menerima hasil analisis kluster, Menafsirkan hasil dari cluster yang dibuat adalah langkah berikutnya, 9) Kesimpulan beserta saran.

HASIL DAN DISKUSI

Hasil penggunaan metode K-Means Clustering terhadap pengolahan data siswa menggunakan aplikasi IBM SPSS Statistic 26 diuraikan sebagai berikut.

Inisialisasi Data

Pada kolom Kabupaten, Asal Sekolah, Jenis Kelamin, Durasi Belajar, Memiliki Laptop, Tempat Tinggal, Memiliki Kendaraan, Beasiswa, Pendapatan Orangtua berupa data kualitatif, karena semua data mempunyai status yang sama dalam artian tidak ada kumpulan data yang lebih unggul atau lebih rendah daripada kumpulan data lainnya, oleh karena itu perlu dilakukan inisialisasi terlebih dahulu data tersebut ke pada bentuk numerik. Tabel berikut menampilkan data yang diinisialisasi peneliti.

Tabel 1. Tabel inisialisasi data

No	Variabel	Inisial	
1	Kabupaten	Kabupaten Nias	1
		Kabupaten Nias Barat	2
		Kabupaten Nias Selatan	3
		Kabupaten Nias Utara	4
		Kota Gunungsitoli	5
		Yang Lain: Kabupaten Tapanuli Tengah	6
2	Asal Sekolah	SMA	1
		SMK	2
		MA	3

3	Durasi Belajar	Kurang dari 1 jam	1
		1 - 2 jam	2
		2 - 3 jam	3
		3 - 4 jam	4
		> 4 Jam	5
4	Memiliki Laptop	Ya	1
		Tidak	2
5	Tempat Tinggal	Kos	1
		Rumah	2
		Asrama	3
		Tempat Saudara	4
6	Memiliki Kendaraan	Ya	1
		Tidak	2
7	Beasiswa	Ya	1
		Tidak	2
8	Pendapatan Orangtua	Kurang dari Rp. 500.000	1
		Rp.500.000 – Rp. 1.000.000	2
		Rp. 1.000.000 – Rp. 2.000.000	3
		Diatas Rp. 2.000.000	4

Data tersebut kemudian disusun kembali berdasarkan Tabel 4.1, dengan definisi variabel operasional sebagai berikut:

- X_1 = Kabupaten/kota Mahasiswa berasal
- X_2 = Asal Sekolah Mahasiswa
- X_3 = Usia mahasiswa
- X_4 = Semester mahasiswa
- X_5 = Rata – rata lama belajar mandiri mahasiswa per hari (Jam)
- X_6 = Jumlah SKS mahasiswa pada semester terakhir
- X_7 = IPS mahasiswa pada semester terakhir
- X_8 = IPK mahasiswa pada semester terakhir
- X_9 = Memiliki Laptop
- X_{10} = Tempat tinggal mahasiswa sekarang
- X_{11} = Memiliki Kendaraan (Sepeda motor)
- X_{12} = Penerima Beasiswa
- X_{13} = Pendapatan Orang Tua Mahasiswa

Uji Asumsi Multikolinearitas

Dalam (Sari & Sukestiyarno, 2021) menyatakan bahwa “Uji multikolinearitas dilakukan untuk menguji ada atau tidaknya variabel bebas yang mempunyai kemiripan antara variabel bebas lain. Besaran (*quality*) yang digunakan untuk mendeteksi adanya multikolinearitas adalah factor inflasi ragam (*Variance Inflation Factor/VIF*)”. Menurut (Musfiani, 2019) “Adanya multikolinearitas adalah suatu pelanggaran dalam analisis cluster karena dapat mempengaruhi hasil cluster”.

Transformasi data merupakan salah satu pendekatan yang dapat digunakan jika data menunjukkan multikolinearitas. Untuk memeriksa multikolinearitas, lakukan hal berikut:

1. Periksa pentingnya toleransi

Jika nilai toleransi kurang dari 0,1 maka terjadi multikolinearitas.

2. Periksa nilai Variance Inflation Factor, atau VIF.

Jika nilai VIF kurang dari 10 maka terjadi multikolinearitas.

Hasil uji asumsi multikolinearitas sebagai berikut:

Tabel 2. Hasil Uji Multikolinearitas

Coefficients ^a			
Model		Collinearity Statistics	
		Tolerance	VIF
1	X1	.811	1.233
	X2	.887	1.128
	X3	.450	2.223
	X4	.239	4.182
	X5	.923	1.084
	X6	.167	5.991
	X7	.103	9.680
	X8	.275	3.639
	X9	.816	1.226
	X10	.772	1.295
	X11	.814	1.228
	X12	.844	1.185
	X13	.828	1.208

Dari Tabel 2. terlihat bahwa “seluruh variabel penelitian tidak mengalami multikolinearitas karena nilai toleransinya lebih besar dari 0,1 dan nilai VIF kurang dari 10, sehingga proses clustering dapat dilanjutkan apabila memenuhi asumsi non-multikolinearitas”.

Standarisasi Data

Standarisasi data sangat diperlukan untuk standarisasi data guna meminimalkan pengaruh perbedaan skala atau satuan. Untuk melakukan ini, data dikurangi dari mean dan dibagi dengan standar deviasi. Standarisasi data dilakukan dengan mengkonversi data ke format z-score menggunakan rumus standarisasi data atau menggunakan IBM SPSS Statistics 26. Berdasarkan hasil yang didapatkan pada statistic deskriptif, nilai standar deviasi $X_1 = 1,468$, $X_2 = 0,494$, $X_3 = 1,444$, $X_4 = 2,136$, $X_5 = 1,019$, $X_6 = 4,393$, $X_7 = 0,63983$, $X_8 = 0,26663$, $X_9 = 0,455$, $X_{10} = 0,697$, $X_{11} = 0,501$, $X_{12} = 0,399$, $X_{13} = 0,989$. Nilai standar deviasi inilah yang akan menentukan nilai z-score.

Asumsi Dalam Analisis Cluster

Sebelum melakukan clustering, harus dipenuhi dua prasyarat yaitu:

Sampel Mewakili Populasi

Nilai KMO dapat digunakan untuk mengetahui apakah sampel tersebut mewakili populasi atau tidak. Jika nilainya lebih besar dari 0,5 maka sampel dianggap representatif; jika kurang dari 0,5, sampel dianggap tidak representatif (Zaki et al., 2022). Nilai KMO dari data ditampilkan dalam tabel berikut. Nilai KMO dari data ditampilkan dalam tabel berikut.

Tabel 4. Nilai KMO

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.624
Bartlett's Test of Sphericity	Approx. Chi-Square	971.870
	Df	78
	Sig.	.000

Berdasarkan Tabel 4. Dapat dikatakan bahwa “nilai KMO sebesar $0,624 > 0,5$, sehingga dapat disimpulkan bahwa sampel tersebut mewakili populasi umum atau merupakan sampel yang representative”.

Nonmultikolinearitas

Berdasarkan hasil uji multikolinearitas pada Tabel 2. terlihat bahwa “seluruh variabel penelitian tidak mengalami multikolinearitas karena nilai toleransinya lebih besar dari 0,1 dan nilai VIF kurang dari 10, sehingga asumsi non multikolinearitas terpenuhi”.

Proses Clustering

Pembentukan K-Means Cluster

Analisis cluster dengan metode *K-means Cluster* dilakukan dengan bantuan IBM SPSS Statistic 26. Pertama ditentukan jumlah *cluster* adalah 4. Hasil Output K-Means Cluster dari SPSS diuraikan sebagai berikut.

Initial Cluster

Presentasi pertama prosedur pengelompokan sebelum proses iterasi. Menurut (Puspita, 2021) “Pada tabel dapat diketahui tampilan awal proses clustering yaitu hasil sementara proses pengelompokan data sebelum dilakukan proses iterasi yang menunjukkan nilai pusat atau centroid cluster awal dari data yang sebelumnya telah dilakukan standarisasi”.

Tabel 5. Centroid Awal Cluster

	Initial Cluster Centers			
	Cluster			
	1	2	3	4
Zscore: X1	-1.79588	-1.79588	.24840	.92983
Zscore: X2	-.67859	-.67859	-.67859	-.67859
Zscore: X3	.78443	2.16913	.09209	-1.29261
Zscore: X4	1.44374	.50738	-1.36532	-1.36532
Zscore: X5	1.70212	-.26112	-1.24274	.72050
Zscore: X6	-1.97464	.75689	.30163	.30163
Zscore: X7	-1.84092	1.23801	-.52808	1.33178
Zscore: X8	-.41460	.86060	-2.25238	2.21081
Zscore: X9	-.63852	-.63852	1.55841	1.55841
Zscore: X10	-.97506	3.32792	1.89359	.45927

Zscore: X11	-1.03255	-1.03255	.96371	-1.03255
Zscore: X12	.49416	-2.01368	.49416	.49416
Zscore: X13	.08468	.08468	.08468	.08468

Iterasi Analisis Cluster

Sepuluh contoh item dimasukkan melalui proses iterasi, seperti terlihat pada tabell.

Tabel 6. Iterasi

Iteration History ^a				
Iteration	Change in Cluster Centers			
	1	2	3	4
1	2.546	2.965	3.268	3.005
2	.559	.763	.467	.342
3	.085	.372	.165	.244
4	.000	.223	.074	.124
5	.000	.241	.000	.108
6	.000	.000	.046	.047
7	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is ,000. The current iteration is 7. The minimum distance between initial centers is 5,971.

Tabel 6. menunjukkan bahwa “proses iterasi dilakukan sebanyak 7 kali. Proses ini dilakukan untuk mendapatkan cluster yang tepat. Terlihat jarak minimum antar pusat cluster yang dihasilkan dari hasil iterasi adalah 5,971”.

Jarak antar cluster

Tabel 7. Jarak Antar Cluster

Distances between Final Cluster Centers				
Cluster	1	2	3	4
1		4.655	4.157	4.253
2	4.655		3.607	2.664
3	4.157	3.607		2.512
4	4.253	2.664	2.512	

Jarak antara centroid dari satu cluster ke cluster lainnya dikenal dengan istilah centroid distance atau jarak antar. Hasil Cluster dianggap baik jika mempunyai nilai jarak yang relatif tinggi. Nilai yang lebih besar berarti jarak antara dua cluster semakin jauh. Berdasarkan Tabel 7, Terdapat 4,655 yang memisahkan cluster 1 dan 2, 4,157 yang memisahkan cluster 1 dan 3, 4,253 yang memisahkan cluster 1 dan 4, yang memisahkan cluster 2 dan 3 dari 3,607 mil yang memisahkan cluster 2 dan 4 dari 2,664, serta jarak antar cluster 3 dan cluster 4 yaitu 2.512. Artinya, perbedaan antara klaster yang satu dengan klaster lainnya semakin terlihat jelas. Dari hasil olah data menggunakan IBM SPSS Statistic 26 didapatkan jarak antar cluster relatif tinggi sehingga dapat dikatakan hasil dari clustering baik.

Jumlah anggota cluster

Tabel berikut menunjukkan jumlah anggota setiap cluster berdasarkan Jumlah Kasus di Setiap

Output Cluster. Dari 203 data mahasiswa, dibentuk 4 buah *cluster* dimana *cluster* 3 merupakan *cluster* dengan jumlah terbanyak yaitu 65 data mahasiswa, 45 data mahasiswa pada *cluster* 1, 29 data mahasiswa pada *cluster* 2, dan 64 data mahasiswa pada *cluster* 4.

Pengujian tingkat signifikan

Uji ANOVA perlu dilakukan untuk melihat uji tingkat signifikansi antar *cluster* dan mengetahui perbedaan tiap *cluster*.. Menurut Hair dalam (Zaki et al., 2022) mengatakan bahwa “untuk menguji apakah ada perbedaan yang signifikan secara statistic antara *cluster* yang terbentuk dengan masing-masing variabel clustering (variabel bebas) adalah dengan menggunakan statistic F dari One-Way-Anova”. Bila menggunakan F hitung dalam analisis klaster, aturannya adalah semakin tinggi F hitung Semakin besar perbedaan antara keempat *cluster* yang dibuat, maka semakin tinggi tingkat signifikansi (sig) < 0,05 dan F hitung (dalam hal pengujian hipotesis lebih besar dari F tabel). Hasil tes ANOVA tercantum di bawah ini.

Tabel 9. Hasil Uji ANOVA

ANOVA						
	Cluster		Error		F	Sig.
	Mean Square	Df	Mean Square	df		
Zscore: X1	14.760	3	.793	199	18.624	.000
Zscore: X2	2.851	3	.972	199	2.932	.035
Zscore: X3	21.912	3	.685	199	32.001	.000
Zscore: X4	44.521	3	.344	199	129.459	.000
Zscore: X5	4.236	3	.951	199	4.453	.005
Zscore: X6	59.206	3	.123	199	483.203	.000
Zscore: X7	56.165	3	.168	199	333.587	.000
Zscore: X8	30.881	3	.550	199	56.194	.000
Zscore: X9	12.286	3	.830	199	14.805	.000
Zscore: X10	14.678	3	.794	199	18.491	.000
Zscore: X11	10.538	3	.856	199	12.308	.000
Zscore: X12	43.241	3	.363	199	119.057	.000
Zscore: X13	4.722	3	.944	199	5.002	.002

Berdasarkan Tabel 9, dapat dilihat bahwa “pada kolom cluster adalah *besaran between cluster means* dan kolom error adalah *besaran within cluster means*”. Sementara itu, rumus berikut menghasilkan kolom F.

$$F = \frac{\text{between cluster means}}{\text{within cluster means}} \quad (1)$$

Hipotesis:

H₀ = Tidak ada variasi yang berarti antara keempat *cluster*.

H₁ = Ada perbedaan mencolok antara keempat *cluster*

Jika tanda. > 0,05, maka H₀ diperbolehkan. H₁ ditolak.

H₀ ditolak jika bilangan sig. <0,05. H₁ disetujui

Variabel yang paling membedakan anggota keempat klaster adalah variabel X6, jumlah SKS, karena mempunyai nilai F paling tinggi diantara variabel-variabel tersebut. Berdasarkan hasil yang

diperoleh seluruh variabel mempunyai nilai signifikansi (sig.) $< 0,05$ yang menunjukkan bahwa keempat cluster mempunyai perbedaan yang signifikan.

Interpretasi *K-Means Cluster*

Pada keluaran akhir pusat klaster dapat dilihat informasi terkait proses standardisasi berdasarkan z-score.

Tabel 10. Hasil Cluster

Final Cluster Centers				
	Cluster			
	1	2	3	4
Zscore: X1	-.05445	-.01007	-.55883	.61041
Zscore: X2	-.27343	.22952	-.11760	.20769
Zscore: X3	.98444	.21146	-.48310	-.29736
Zscore: X4	1.44374	.08764	-.70267	-.34119
Zscore: X5	.43692	-.22727	-.20071	-.00038
Zscore: X6	-1.74701	.54496	.39619	.57906
Zscore: X7	-1.62038	.86453	.08218	.66413
Zscore: X8	-.49127	.98864	-.68752	.59572
Zscore: X9	-.54088	-.18398	.57824	-.12361
Zscore: X10	-.01884	.55819	-.62199	.39203
Zscore: X11	-.01224	-.48185	.53375	-.31514
Zscore: X12	.10405	-1.92721	.30124	.49416
Zscore: X13	.24196	-.08966	-.35089	.22687

Tabel 10. menunjukkan hasil proses clustering akhir yang menghasilkan 4 cluster untuk setiap variabel. Nilai-nilai standar menghasilkan variabel-variabel yang membentuk tabel pusat cluster akhir. Jika angkanya negatif, berarti data tersebut berada di bawah rata-rata keseluruhan; jika positif berarti datanya di atas rata-rata (Sari & Sukestiyarno, 2021). Berdasarkan Tabel 10. didapatkan karakteristik pada masing-masing cluster ialah

1. *Cluster 1* berisi variabel X₃, X₄, X₅, X₁₂, X₁₃, berada di atas rata-rata total serta X₁, X₂, X₆, X₇, X₈, X₉, X₁₀, X₁₁ berada di bawah rata-rata total.
2. *Cluster 2* berisi variabel X₂, X₃, X₄, X₆, X₇, X₈, X₁₀ berada di atas rata-rata total serta X₁, X₅, X₉, X₁₁, X₁₂, X₁₃ berada di bawah rata-rata total.
3. *Cluster 3* berisi variabel X₆, X₇, X₉, X₁₁, X₁₂, berada di atas rata-rata total serta X₁, X₂, X₃, X₄, X₅, X₈, X₁₀, X₁₃ berada di bawah rata-rata total.
4. *Cluster 4* berisi variabel X₁, X₂, X₆, X₇, X₈, X₁₀, X₁₂, X₁₃ berada di atas rata-rata total serta X₃, X₄, X₅, X₉, X₁₁ berada di bawah rata-rata total.

Dimana Statistik deskriptif hasil cluster yang diperoleh ialah *Cluster 1* untuk variabel Kabupaten terdiri atas 7 orang Kabupaten Nias, 4 orang Kabupaten Nias Barat, 9 orang Kabupaten Nias Selatan, 7 orang Kabupaten Nias Utara, dan 18 orang di Kota Gunungsitoli. Variabel asal sekolah terdiri atas 33 orang SMA dan 12 orang SMK. Rata-rata usia mahasiswa di cluster 1 adalah 22 tahun dan hanya diisi oleh semester 8 sebanyak 45 orang. Dengan rata-rata durasi belajar 18 orang kurang dari jam, 12 orang 1-2 jam, 5 orang 2-3 jam, 9 orang 3-4 jam serta 1 orang lebih dari 4 jam.

Jumlah rata-rata SKS adalah 13 dengan IPS rata-rata 1,77 dan IPK rata-rata 2,98. Yang memiliki laptop 43 orang dan yang tidak 2 orang. Tempat tinggal kos sebanyak 20 orang, rumah sebanyak 21 orang, asrama 1 orang serta 3 orang di tempat saudara. Terdapat 22 orang yang memiliki mobil, khususnya sepeda motor, dan 23 orang tidak. Ada 38 orang yang tidak menerima beasiswa dan 7 orang yang menerima. Orang tuanya berpenghasilan antara Rp. 500.000 dan Rp. 1.000.000 untuk 12 orang, Rp. 500.000 dan Rp. 2.000.000 untuk 5 orang, dan Rp. 2.000.000 ke atas untuk 10 siswa. Penghasilan orang tuanya kurang dari Rp. 500.000 untuk 18 orang..

Cluster 2 untuk variabel Kabupaten terdiri atas 7 orang Kabupaten Nias, 1 orang Kabupaten Nias Barat, 1 orang Kabupaten Nias Selatan, 7 orang Kabupaten Nias Utara, dan 13 orang di Kota Gunungsitoli. Variabel asal sekolah terdiri atas 15 orang SMA dan 14 orang SMK. Rata-rata usia mahasiswa di cluster 2 adalah 21 tahun dan diisi oleh Ada tiga orang pada semester kedua, tujuh orang pada semester keempat, dan sembilan belas orang pada semester enam. 9 peserta menyelesaikan survei dalam waktu kurang dari satu jam, 12 dalam 1-2 jam, 6 dalam 2-3 jam, dan rata-rata 2 dalam 3-4 jam. Dengan rata-rata IPS 3,40 dan rata-rata IPK 3,37, rata-rata beban kredit adalah 23. Yang memiliki laptop 23 orang dan yang tidak 6 orang. Tempat tinggal kos sebanyak 6 orang, rumah sebanyak 17 orang, asrama 2 orang serta 4 orang di tempat saudara. Terdapat 21 orang yang memiliki mobil khususnya sepeda motor dan 8 orang tidak. Satu orang tidak mendapatkan beasiswa, sedangkan penerima hibah ada 28 orang. Kurang dari Rp. 500.000 adalah penghasilan orang tuanya untuk 12 orang, antara Rp. 500.000 dan Rp. 1.000.000 untuk 11 orang, antara Rp. 1.000.000 dan Rp. 2.000.000 untuk 5 orang, dan di atas Rp. 2.000.000 untuk 1 siswa..

Cluster 3 untuk variabel Kabupaten terdiri atas 13 orang Kabupaten Nias, 15 orang Kabupaten Nias Barat, 13 orang Kabupaten Nias Selatan, 19 orang Kabupaten Nias Utara, dan 5 orang di Kota Gunungsitoli. Variabel asal sekolah terdiri atas 45 orang SMA dan 20 orang SMK. Rata-rata usia mahasiswa di cluster 3 adalah 20 tahun dan diisi oleh Semester kedua berjumlah 28 orang, semester empat 28 orang, dan semester enam 9 orang. 21 peserta menyelesaikan penelitian dalam waktu kurang dari satu jam, 26 dalam 1-2 jam, 11 dalam 2-3 jam, dan rata-rata 7 dalam 3-4 jam. Dengan rata-rata IPK 2,93 dan rata-rata IPS 2,90, rata-rata terdapat 22 SKS. Yang memiliki laptop 29 orang dan yang tidak 36 orang. Tempat tinggal kos sebanyak 52 orang, rumah sebanyak 10 orang serta 3 orang di tempat saudara. Yang memiliki kendaraan khususnya sepeda motor 14 orang dan yang tidak sebanyak 51 orang. Ada 60 orang yang tidak menerima beasiswa dan 5 orang yang menerima. Orang tuanya berpenghasilan antara Rp. 500.000 dan Rp. 1.000.000 untuk 25 orang, Rp. 500.000 dan Rp. 2.000.000 untuk 3 orang, dan Rp. 2.000.000 ke atas untuk 2 orang murid. Penghasilan orang tuanya kurang dari Rp. 500.000 untuk 35 orang..

Cluster 4 untuk variabel Kabupaten terdiri atas 2 orang Kabupaten Nias, 3 orang Kabupaten Nias Barat, 1 orang Kabupaten Nias Selatan, 12 orang Kabupaten Nias Utara, 45 orang di Kota Gunungsitoli dan 1 orang di luar yaitu Kabupaten Tapanuli Tengah. Variabel asal sekolah terdiri atas 38 orang SMA dan 24 orang SMK. Mahasiswa semester 2 berjumlah 14 orang, semester 4 berjumlah

30 orang, dan semester 6 berjumlah 20 orang. Rata-rata usia mahasiswa gugus 4 adalah 20 tahun. Dua belas peserta menyelesaikan penelitian dalam waktu kurang dari satu jam, dua puluh tujuh dalam 1-2 jam, dua puluh dua dalam 2-3 jam, dua dalam 3-4 jam, dan rata-rata satu dalam lebih dari 4 jam. Dengan rata-rata IPS 3,27 dan rata-rata IPK 3,27, rata-rata terdapat 23 SKS.. Yang memiliki laptop 49 orang dan yang tidak 15 orang. Tempat tinggal kos sebanyak 11 orang, rumah sebanyak 45 orang, serta 8 orang di tempat saudara. Yang memiliki kendaraan khususnya sepeda motor 41 orang dan yang tidak sebanyak 23 orang. Penerima beasiswa tidak ada serta yang tidak sebanyak 64 orang. Kurang dari Rp. 500.000 untuk 21 orang, antara Rp. 500.000 dan Rp. 1.000.000 untuk 22 orang, antara Rp. 1.000.000 dan Rp. 2.000.000 untuk 12 orang, dan lebih dari Rp. 2.000.000 untuk maksimal 9 murid adalah penghasilan orangtuanya.

KESIMPULAN

Kesimpulan yang dapat diambil dari hasil dan perdebatan yang telah diuraikan adalah 1) Penerapan metode cluster analysis dalam pembentukan cluster pada data mahasiswa program studi pendidikan matematika menghasilkan cluster yang baik dimana tiap cluster memiliki perbedaan signifikan dibuktikan dengan nilai uji ANOVA signifikan < 0.05 dan faktor-faktor yang paling penting dalam pembentukan cluster diperiksa satu demi satu dari nilai F adalah Jumlah SKS, IPS, Semester, Beasiswa, IPK, Usia, Kabupaten, Tempat Tinggal, Memiliki Laptop, Memiliki Kendaraan, Pendapatan Orangtua, Rata-rata durasi belajar dan yang terakhir Asal Sekolah. 2) Cluster yang terbentuk dalam proses cluster analysis dengan menggunakan K-Means Cluster Method terdiri atas 4 cluster dengan jumlah tiap anggota cluster ialah Catatan siswa di Klaster 1 berjumlah 45 orang, Klaster 2 berjumlah 29 orang, dan Klaster 3 berjumlah 65 orang, serta cluster 4 sebanyak 64 data mahasiswa. 3) Hasil cluster analisis didapatkan karakteristik pada masing-masing cluster ialah cluster 1 berisi variabel usia, semester, rata-rata durasi belajar, beasiswa, pendapatan orang tua, di atas rata-rata keseluruhan serta kabupaten, asal sekolah, jumlah SKS, IPS, IPK, memiliki laptop, tempat tinggal, memiliki kendaraan berada di bawah rata-rata total. cluster 2 berisi variabel asal sekolah, usia, semester, jumlah SKS, IPS, IPK, tempat tinggal berada di atas rata-rata total serta kabupaten, rata-rata durasi belajar, memiliki laptop, memiliki kendaraan, beasiswa, pendapatan Rata-rata orang tua lebih rendah dibandingkan kelompok. Variabel angka terdapat pada cluster 3 SKS, IPS, memiliki laptop, memiliki kendaraan, beasiswa, berada di atas rata-rata total serta kabupaten, asal sekolah, usia, semester, rata-rata durasi belajar, IPK, tempat tinggal, pendapatan orang tua kurang dari rata-rata keseluruhan. Dan cluster 4 berisi variabel kabupaten, asal sekolah, jumlah SKS, IPS, IPK, tempat tinggal, beasiswa, pendapatan orang tua di atas rata-rata keseluruhan serta usia, semester, rata-rata durasi belajar, memiliki laptop, memiliki kendaraan berada di bawah rata-rata keseluruhan. 4) Mahasiswa pendidikan matematika paling banyak berasal dari Kota Gunungsitoli dan latar belakang berasal dari SMA. Mahasiswa pada semester 8 semuanya dibandingkan dengan mahasiswa pada semester sebelumnya, termasuk dalam cluster yang sama. Hal ini disebabkan

pengaruh jumlah SKS yang diambil. Semester 8 cenderung mengambil jumlah SKS sedikit dibandingkan mahasiswa semester lainnya. Mahasiswa dengan IPK yang tinggi masuk ke Cluster 2 dengan rata-rata 3.37. Mahasiswa dengan IPK yang sedang masuk ke Cluster 4 dengan rata-rata 3.27, Mahasiswa dengan IPK yang rendah masuk ke Cluster 1 dengan rata-rata 2.98, dan Mahasiswa dengan IPK yang terendah masuk ke dalam Cluster 3 dan memiliki rata-rata 2,93.

UCAPAN TERIMA KASIH

Segala puji dan syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena telah membuat kegiatan ini berjalan dengan baik. Saya ingin mengucapkan terima kasih kepada semua orang tua, dosen pembimbing, dan rekan mahasiswa terkait yang telah mendukung dan menyemangati saya—saya tidak bisa berterima kasih kepada mereka semua secara individu.

REFERENSI

- awaliah, R. (2018). *Analisis Clustering Untuk Mengelompokkan Tingkat Kesejahteraan Kabupaten/Kota Berdasarkan Sosial Ekonomi Rumah Tangga Di Wilayah Provinsi Sulawesi Selatan* [Universitas Islam Negeri (Uin) Alauddin Makassar]. [Http://Repositori.Uin-Alauddin.Ac.Id/Id/Eprint/8983](http://Repositori.Uin-Alauddin.Ac.Id/Id/Eprint/8983)
- Ayuningtias, Indah, Debataraja, Naomi Nessyana, & Imro'ah, Nurfitri. (2019). Analisis Cluster Non-Hirarki Dengan Metode K-Modes. *Buletin Ilmiah Mat. Stat. Dan Terapannya (Bimaster)*, Volume 08, No. 4, 909-916.
- Lestari, A. D. (2018). *Analisis Multivariat Clustering K-Means Pada Kabupaten/Kota Di Provinsi Jawa Tengah Berdasarkan Indeks Pembangunan Manusia Tahun 2017 Dengan Bantuan Software Spss*. [Http://Lib.Unnes.Ac.Id/Id/Eprint/36913](http://Lib.Unnes.Ac.Id/Id/Eprint/36913)
- Lestari, Whidi. (2019). Clustering Data Mahasiswa Menggunakan Algoritma K-Means Untuk Menunjang Strategi Promosi (Studi Kasus : Stmik Bina Bangsa Kendari). *Jurnal Sistem Informasi Dan Sistem Komputer*, Vol. 4, No. 2. [Http://E-Jurnal.Stmikbinsa.Ac.Id/Index.Php/Simkom](http://E-Jurnal.Stmikbinsa.Ac.Id/Index.Php/Simkom)
- Marajari, M. R., Nasution, K., & Haramaini, T. (2021). Analisis Data Mahasiswa Untuk Mendukung Strategi Promosi Pada Universitas Medan Area Menggunakan Algoritma K-Means Clustering. *Semnastek*, 100–107. [Https://Jurnal.Uisu.Ac.Id/Index.Php/Semnastek/Article/View/4150](https://Jurnal.Uisu.Ac.Id/Index.Php/Semnastek/Article/View/4150)
- Maukar, Anastasia L., Marisa, Fitri, Widodo, Anang Aris, Kamilaningtyas, Nurin, & Nugraha, Novian Didik. (2022). Analisis Data Penerimaan Mahasiswa Baru Berbasis K-Means. *Jiko (Jurnal Informatika Dan Komputer)*, Vol. 6, No. 2, 142–147.
- Muliono, Rizki, & Sembiring, Zulfikar. (2019). Data Mining Clustering Menggunakan Algoritma K-Means Untuk Klasterisasi Tingkat Tridarma Pengajaran Dosen. *Cess (Journal Of Computer Engineering System And Science)*, Vol. 4 No. 2
- Musfiani. (2019). Analisis Cluster Dengan Menggunakan Metode Partisi Pada Pengguna Alat

- Kontrasepsi Di Kalimantan Barat. *Bimaster : Buletin Ilmiah Matematika, Statistika Dan Terapannya*, 8(4), 893–902. <https://doi.org/10.26418/bimst.v8i4.36584>
- Nasari, Fina, & Darma, Surya. (2015). Penerapan K-Means Clustering Pada Data Penerimaan Mahasiswa Baru (Studi Kasus : Universitas Potensi Utama). *Stmik Amikom*, 73-78.
- Nisa, K. (2019). *Analisis Cluster Dengan Menggunakan Metode Hierarki Untuk Pengelompokan Kecamatan Di Kabupaten Langkat Berdasarkan Indikator Kesehatan* [Universitas Islam Negeri Sumatera Utara]. <http://repository.uinsu.ac.id/eprint/11406>
- Nugroho, Nursatio, & Adhinata, Faisal Dharma. (2022). Penggunaan Metode K-Means Dan K-Means++ Sebagai Clustering Data Covid-19 Di Pulau Jawa. *Teknika*, Volume 11(3), 170-179.
- Parlambang, Bagas, & Fauziah. (2020). Implementasi Algoritma K-Means Dalam Proses Penilaian Kuesioner Kepada Dosen Guna Mendukung Kepuasan Mahasiswa Terhadap Dosen. *Jurnal Ilmiah Teknologi Dan Rekayasa*, Vol 25 No. 2. <https://doi.org/10.35760/tr.2020.v25i2.2719>
- Puspita, R. N. (2021). Analisis K-Means Cluster Pada Kabupaten/Kota Di Provinsi Banten Berdasarkan Indikator Indeks Pembangunan Manusia. *Jurnal Lebesgue : Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, 2(3), 267–281. <https://doi.org/10.46306/lb.v2i3.85>
- Puntoriza, Fibriani, Charitas. (2020). Analisis Persebaran Umkm Kota Malang Menggunakan Cluster K-Means. *Journal Of Information System*, Vol. 5, No. 1, 86-94. Doi: 10.33633/joins.v5i1.3469.
- Riana, D. S. (2021). *Analisis Cluster Untuk Mengklasifikasi Tingkat Kesejahteraan Sosial Masyarakat Di Kabupaten Deli Serdang Menggunakan Fuzzy C-Mean Clustering Saat Pandemi Covid-19* [Universitas Islam Negeri Sumatera Utara]. http://repository.uinsu.ac.id/view/creators/riana=3adwi_salasa=3a=3a.default.html
- Rokhimah, W., & Kismiantini. (2022). *Analisis Clustering Tingkat Kerawanan Wilayah Terhadap Kasus Penyakit Di Kabupaten Sleman Dengan Metode K-Means*. 8, 25–37. <https://journal.student.uny.ac.id/ojs/index.php/jktm/article/view/18202>
- Sari, D. N. P., & Sukestiyarno, Y. L. (2021). *Analisis Cluster Dengan Metode K-Means Pada Persebaran Kasus Covid-19 Berdasarkan Provinsi Di Indonesia*. 4, 602–610. <https://journal.unnes.ac.id/sju/index.php/prisma/article/view/45035>
- S.Y., Hasyrif, Rismayani, & Syam, Asrul. (2019). Data Mining Menggunakan Algoritma K-Means Pengelompokan Penyebaran Diare Di Kota Makassar. *Prosiding Seminar Ilmiah Sistem Informasi Dan Teknologi Informasi*, Vol. Viii, No. 1, 73 – 82.
- Zaki, A., Irwan, & Sembe, I. A. (2022). Penerapan K-Means Clustering Dalam Pengelompokan Data (Studi Kasus Profil Mahasiswa Matematika Fmipa Unm). *Jmathcos (Journal Of Mathematics, Computations, And Statistics)*, 5(2), 163–176. <https://doi.org/10.35580/jmathcos.v5i2.38820>
- Zega, Y. (2022). Pengembangan Modul Pembelajaran Berbasis Etnomatematika Pada Materi

Lingkar. In *Jssa: Journal Of Smart Society Adpertis* (Vol. 1, Issue 1).
<https://jurnal.adpertisi.or.id/index.php/jssa/article/view/259>