

Clustering Menggunakan K-Means Untuk Menentukan Mahasiswa Berprestasi (MAWAPRES)

Ghufron¹, Dedy Kurniadi², Andre Sugiyono³

^{1,2} Teknik Informatika, Fakultas Teknologi Industri, Universitas Islam Sultan Agung

³ Teknik Industri, Fakultas Teknologi Industri, Universitas Islam Sultan Agung

e-mail: ghufron@unissula.ac.id, ddy.kurniadi@unissula.ac.id, andre@unissula.ac.id

Abstrak

To analyze and group data into the same groups, clustering is an effective technique. The aim of this study is to identify groups of students who perform based on IPK data, non-academic performance, organizational participation, and foreign language skills. Furthermore, the K-Means clustering algorithm is used to analyze the data of the classification item. Determining the ideal number of groups or centroids is a clustering process. To group out successful and well-performed students, the most suitable group is selected based on the resulting intra-group and intergroup values. It is expected that the results of this study will provide a better understanding of the characteristics of the student group that performs. Educational institutions can use this data to find potential students and provide appropriate support to improve the quality of education. The research concluded that the K-Means clustering method could be used as an effective way to group students based on their academic achievements. This method can help educational institutions find students who perform to be followed in national and international mawapres.

Kata Kunci : Identifikasi, Mawapres, Clustering of K-Means

Abstract

Untuk menganalisis dan mengelompokkan data ke dalam kelompok yang sama, clustering adalah teknik yang efektif. Tujuan dari penelitian ini adalah untuk mengidentifikasi kelompok mahasiswa yang berprestasi berdasarkan data IPK, prestasi non-akademik, keikutsertaan organisasi, dan kemampuan bahasa asing. Selanjutnya, algoritma clustering K-Means digunakan untuk menganalisis data item penilaian. Menentukan jumlah kelompok atau centroid yang ideal adalah proses clustering. Untuk mengelompokkan mahasiswa yang berprestasi dan cukup berprestasi, kelompok yang paling sesuai dipilih berdasarkan nilai intra-kelompok dan inter-kelompok yang dihasilkan. Diharapkan bahwa hasil penelitian ini akan memberikan pemahaman yang lebih baik tentang ciri-ciri kelompok mahasiswa yang berprestasi. Institusi pendidikan dapat menggunakan data ini untuk menemukan mahasiswa yang berpotensi dan memberikan dukungan yang tepat untuk meningkatkan kualitas pendidikan. Penelitian ini menyimpulkan bahwa metode clustering K-Means dapat digunakan sebagai cara yang efektif untuk mengelompokkan siswa berdasarkan prestasi akademik mereka. Metode ini dapat membantu lembaga pendidikan menemukan mahasiswa yang berprestasi untuk diikutkan dalam mawapres nasional dan internasional.

Keywords: *Identifikasi, Mawapres, Clustering K-Means*

1. PENDAHULUAN

Dalam dunia pendidikan tinggi, mengenali dan mendukung mahasiswa berprestasi (MAWAPRES) merupakan hal yang penting. Identifikasi mahasiswa berprestasi dapat membantu institusi pendidikan mengakui dan memberikan penghargaan atas prestasi mereka, serta memberikan dukungan yang sesuai untuk pengembangan mereka di bidang akademik. Namun, dengan jumlah mahasiswa yang besar, tugas ini dapat menjadi rumit dan memakan waktu[1].

Clustering atau pengelompokan merupakan salah satu metode yang dapat digunakan untuk mengelompokkan data menjadi kelompok-kelompok yang serupa berdasarkan karakteristik tertentu[2]. Dalam konteks ini, penelitian ini bertujuan untuk menggunakan metode clustering K-Means sebagai alat untuk menentukan kelompok[3] mahasiswa berprestasi berdasarkan data akademik[4].

Metode K-Means adalah salah satu teknik clustering yang populer dan sering digunakan dalam analisis data[5]. Metode ini bekerja dengan mempartisi data menjadi k kelompok yang saling eksklusif berdasarkan jarak antara titik data dengan centroid kelompok. Dalam konteks penelitian ini, setiap mahasiswa akan diberikan atribut berdasarkan prestasi akademik mereka, seperti ipk, prestasi akademik, dan sebagainya. Data ini akan digunakan untuk menentukan kelompok atau cluster yang sesuai dengan karakteristik prestasi akademik.

Penelitian ini memiliki potensi untuk memberikan wawasan yang berharga tentang karakteristik kelompok mahasiswa berprestasi. Dengan menggunakan metode clustering K-Means, diharapkan dapat diidentifikasi kelompok mahasiswa dengan pola prestasi akademik yang serupa.

Informasi ini dapat digunakan oleh institusi pendidikan untuk mengenali mahasiswa berprestasi dan memberikan dukungan yang tepat, seperti program pendidikan khusus atau pengembangan potensi.

Dengan pemahaman yang lebih baik tentang mahasiswa berprestasi, institusi pendidikan dapat mengoptimalkan upaya mereka dalam meningkatkan kualitas pendidikan, memberikan penghargaan yang sesuai, serta menciptakan lingkungan yang kondusif bagi pengembangan potensi akademik mahasiswa[6].

Dalam penelitian ini, kami akan menggambarkan langkah-langkah yang diambil untuk menerapkan metode clustering K-Means[7] dalam menentukan kelompok mahasiswa berprestasi berdasarkan data akademik mereka. Selanjutnya, kami akan menganalisis hasil clustering dan memberikan pemahaman yang lebih dalam tentang karakteristik kelompok mahasiswa berprestasi.

2. METODOLOGI PENELITIAN

1. Pengumpulan Data

Mengumpulkan data akademik mahasiswa yang relevan, seperti nilai rata-rata, jumlah mata kuliah yang diambil, jumlah skor prestasi lainnya, dan atribut lain yang relevan dalam menentukan prestasi akademik mahasiswa.

2. Pra-pemrosesan Data

Melakukan pra-pemrosesan data untuk membersihkan dan mempersiapkan data agar sesuai dengan analisis clustering. Mengidentifikasi dan mengatasi nilai yang hilang atau tidak valid. Menstandarisasi data jika diperlukan, seperti normalisasi atau penskalaan.

3. Pemilihan Jumlah Cluster

Melakukan pemilihan jumlah cluster yang optimal untuk mempartisi data menggunakan metode K-Means. Menggunakan metode evaluasi internal *davies bouldin index* untuk membantu menentukan jumlah cluster yang paling sesuai[8].

4. Penerapan K-Means Clustering

Mengimplementasikan algoritma K-Means clustering untuk mempartisi data menjadi kelompok berdasarkan atribut prestasi akademik mahasiswa. Menentukan centroid awal secara acak atau menggunakan metode inisialisasi yang lebih canggih, seperti K-Means++. Mengiterasi proses clustering hingga konvergensi, di mana posisi centroid tidak berubah secara signifikan atau jumlah iterasi yang ditentukan tercapai.

Langkah sebagai berikut

1. Tentukan centroid terlebih dahulu
2. Hitung jarak dengan *Euclidean Distance*[9]

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1)$$

5. Evaluasi Hasil Clustering

Mengevaluasi hasil clustering dengan mengukur kualitas partisi yang dihasilkan. Menggunakan metrik evaluasi clustering *Davies Bouldin Index*[10]. Menganalisis hasil clustering untuk memahami karakteristik dan pola prestasi akademik pada setiap kelompok mahasiswa.

6. Interpretasi Hasil

Menganalisis karakteristik dan pola prestasi akademik dalam setiap kelompok mahasiswa berprestasi yang dihasilkan oleh clustering. Mengidentifikasi atribut penting atau faktor yang mempengaruhi klasifikasi mahasiswa sebagai mahasiswa berprestasi. Memberikan interpretasi dan pemahaman yang mendalam tentang karakteristik kelompok mahasiswa berprestasi berdasarkan hasil clustering.

7. Validasi dan Verifikasi

Menggunakan teknik validasi, seperti validasi silang (cross-validation), untuk memastikan keandalan dan kestabilan hasil clustering. Memverifikasi temuan dengan melakukan analisis statistik atau melibatkan ahli domain dalam mengevaluasi hasil clustering[11]. Metode clustering K-Means yang diterapkan dalam penelitian ini akan membantu mengelompokkan

mahasiswa berprestasi berdasarkan atribut prestasi akademik mereka. Dengan mengikuti metodologi ini, diharapkan dapat memberikan wawasan yang berharga tentang karakteristik kelompok mahasiswa berprestasi, yang dapat digunakan oleh institusi pendidikan untuk mengenali dan memberikan dukungan yang tepat kepada mereka.

3. HASIL DAN PEMBAHASAN

Penelitian ini data yang digunakan berasal dari data mahasiswa teknik informatika tahun 2022, sedangkan data yang digunakan adalah ipk mahasiswa, prestasi non akademik, keikutsertaan organisasi dan ketrampilan bahasa asing, proses clustering ini dilakukan untuk mengetahui cluster mahasiswa yang memenuhi sebagai mawapres.

Tabel 1. Data Mahasiswa beserta atribut

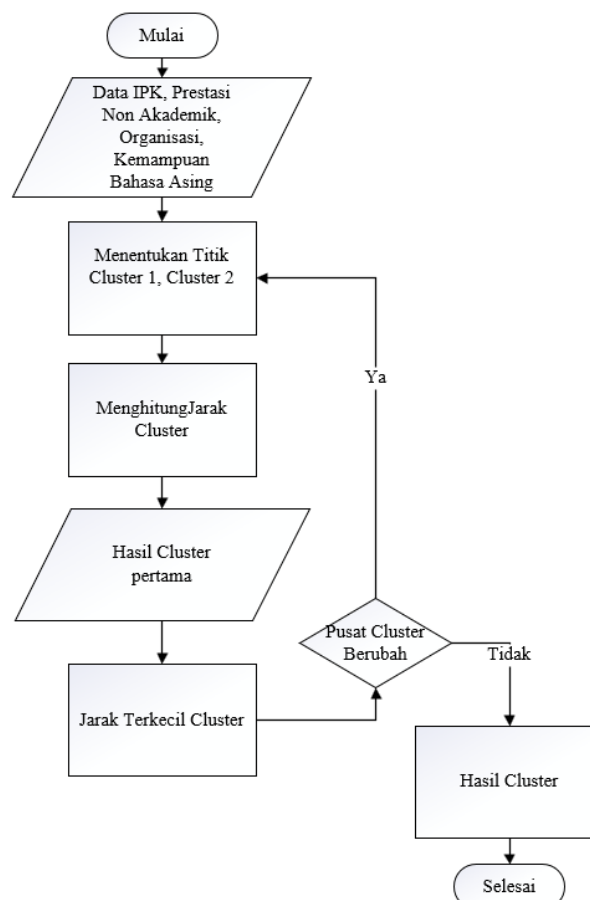
<i>Nama</i>	<i>IPK</i>	<i>Non Akademik</i>	<i>Ikut Organisasi</i>	<i>Kemampuan Bahasa</i>
Saeful Anam	2.8	4	1	74
Muhammad Luthfi	3.4	3	0	84
Widodo	2.6	1	0	79
Taufik Febriansah	3	2	1	82
Silvi Nurcahya	3.3	3	1	81
Aninda	2.3	4	0	76
Abdullah Salim	3.8	2	1	79
Farrel Iqbal	2.9	3	0	78
Ainun Nizar	2.7	3	0	77
Arfiana	3.7	4	1	85
Iven Aurananta	3.3	3	1	87
Nurhasan	3.3	2	1	89
Supriyatin	3.3	1	1	91
Setianingrum	3.3	3	0	93
Muhammad Fatih	3.5	4	1	95
Royhan	3.3	3	1	97
Arfiyanto Fiko	2.7	2	1	80
Tatik Lusiana	2.1	4	0	76
Qomaruz Zaman	2.6	3	0	75
Febrianto	3.1	3	1	83
Muhamad Azis	3.5	3	0	84
Abdul Latif	2.8	2	0	76
Khusna Faizuddin	3.4	3	1	86
Achmad Roychan	3.4	2	1	87
Adhistry Auliatuzzahra	3.3	2	1	88
Aditya Bachtiar	3.5	2	1	89
Aditya Sendy	3.6	2	1	90
Afsha Alifia	3.2	3	1	91
Agung Widy	2.8	3	1	92
Ahmad Fahim	3.6	3	1	93
Ahmad Faiz	3	2	1	94
Ahmad Rafi	3.1	2	1	95
Aisa Putri Maryam	3.6	3	1	96
Akbar Dzidan	3.5	3	1	97
Alessandra Trishta	3.3	3	1	91
Alfida	3.4	4	1	89
Alya	3.3	3	1	87

<i>Nama</i>	<i>IPK</i>	<i>Non Akademik</i>	<i>Ikut Organisasi</i>	<i>Kemampuan Bahasa</i>
Andika Arif	3.1	3	1	85
Anindya Farah	3.3	3	1	83
Ardan Chaliya	3.4	3	1	81
Arjun Khoirul Abror	3.3	3	1	79
Asep Hernandi	3.4	3	1	77
Assyfa Febriwanti	3.3	3	1	75
Avenia Rohmatun Nisa	3.4	3	1	73
Benaya Maula Syafiq	3.1	3	0	78
Dafavico Assechan	2.9	3	0	69
Daililah Syahrina Aulia	2.7	3	1	69
Damar Riyadi Syahputra	3	3	2	76
David Dimas Santana	3.3	3	2	78
Dian Purbo Wicaksono	2.4	0	0	75
Difka Fannansyah	3.1	2	1	78
Dimas Taufikurrohman	2.7	2	0	74
Diva Aequo Putri Laila Ali	2.8	2	1	75
Dwiky Chandra Mulyo Utomo	3.3	2	1	84
Ella Heriyawati	3.4	3	1	82
Eva Sopitri	3.1	2	1	79
Fajar Octanouva	2.9	2	1	79
Fandy Ahmad Rizal	3	2	1	80
Farhan Nur Hidayat	3.3	2	1	81
Fazya Tiara Artamevia	3.3	2	1	82
Febriana Susilowati	3.3	2	1	83
Fedrick Rakapati	3.2	2	1	84
Franklyn Rama Fitrah Akbar	2.9	2	0	68
Gilang Setiawan	3	2	1	87
Gusmix Mubarak	2.6	2	0	67
Habib Baihaqi Arrozaq	3	2	1	97
Hadana Maulana	3.1	2	1	80
Hanikun Sukma Salsabiella	3.1	2	1	82
Hawa Parida Nurlita	2.9	2	1	84
Helmi Septianto	3	2	2	86
Heru Hidayatulloh	2.9	2	1	88
Ibra Andnan Faiz Sempana	3	2	1	90
Ikharista Ayu Nusrotun A	3.2	2	1	92
Ikrana Kharisa Hanum	3.1	2	1	90
Jibril Amjad Taqiudin	2.6	2	0	81
Keisya Anazwa	3.1	2	1	98
Luluk Indah	3.3	2	1	83
Lutfi Febrimaharani	1.2	2	0	68
Manar Nuha Afifah	3.4	2	1	80
Ma Ruf	2	2	0	77
Maulana Alfi	2.6	2	1	75
Maulana Arlistyo	3	2	1	83
Maulana Farid	2.8	2	0	71
Maulana Ishak	2.7	2	1	78
Mohammad Riyan	2.7	2	0	76

<i>Nama</i>	<i>IPK</i>	<i>Non Akademik</i>	<i>Ikut Organisasi</i>	<i>Kemampuan Bahasa</i>
Moh. Zidni Iman	3.3	2	0	80
M. Sirajuth Thayyib	3.1	2	0	79
Muchammad Harun	1.7	2	0	78
Muh. Akhyarul Insan	3.3	2	0	84
Muhamad Huda	2.7	2	0	69
Muhamad Pagah	3.1	2	0	89
Muhammad Zainurrohim	3.2	2	1	82
Muhammad Aris	2.2	2	0	78
Muhammad Azizan	3.5	2	1	79
Muhammad Faishal	2.7	2	0	78
Muhammad Fauzan	3.5	2	1	75
Nur Arifin	3.3	2	1	78
Muhammad Ridwan	3.9	4	1	91

3.1. Proses Clustering dengan K-Means

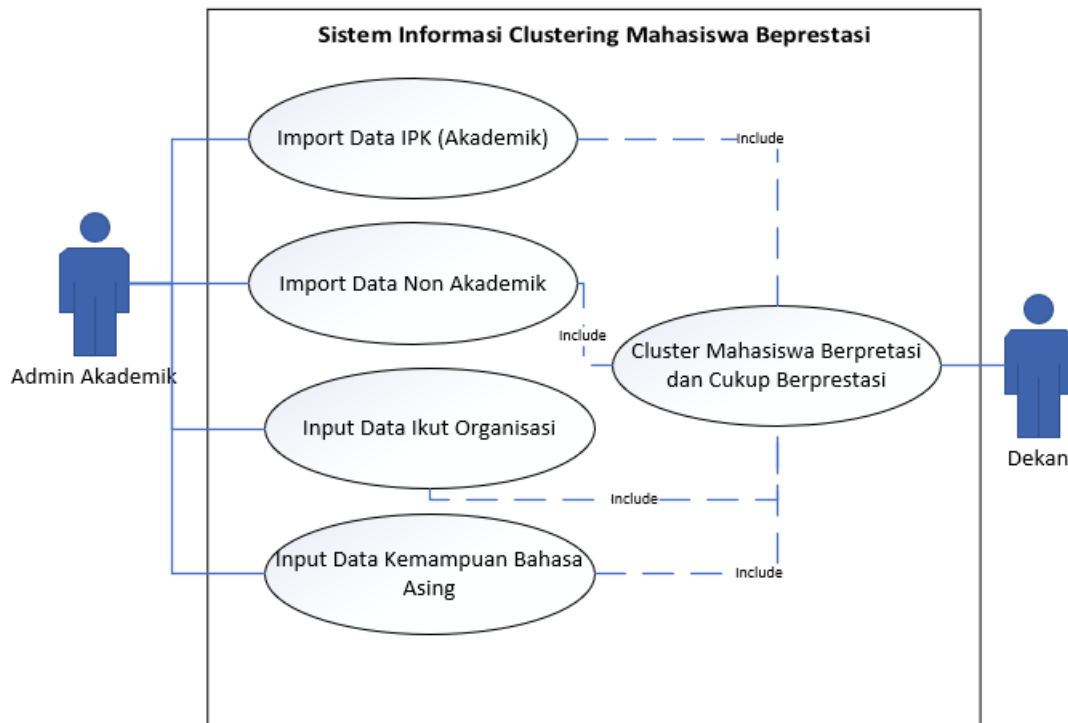
Berdasarkan data-data pendukung mahasiswa yaitu prestasi akademik, prestasi non akademik, ikut serta organisasi dan kemampuan bahasa asing yang digunakan untuk melakukan proses clustering dengan k-means, Proses clustering ini dilakukan dengan cara menentukan titik cluster atau centroid setelah itu dihitung untuk mencari jarak dari setiap cluster, ketika jarak cluster telah ditemukan, hasil cluster dikelompokkan berdasarkan jarak dari setiap cluster tersebut. Apabila terdapat hasil iterasi titik dari cluster satu dengan cluster yang lain, maka dilakukan perhitungan cluster kembali sampai titik cluster tidak berubah. proses alur clustering Dapat dilihat pada gambar 2. flowchart K-means Clustering untuk menentukan mahasiswa berprestasi.



Gambar 1. Flowchart Proses Sistem Clustering K-Means

3.2. Implementasi Clustering K-Means

Untuk memudahkan analisa dan implementasi pada pengembangan sistem ini dibuatlah *use case* diagram untuk menunjukkan hubungan antara admin akademik dan dekan, admin akademik bertugas mengimport data dari sistem akademik dengan koneksi *application programming interface* (API) dan import data prestasi non akademik dari sistem kegiatan mahasiswa lewat API dengan cara klik data, kemudian untuk input data ikut organisasi dan kemampuan berbahasa di inputkan secara langsung kedalam sistem lebih detail dapat dilihat pada gambar use case berikut



Gambar 2. Usecase Sistem Informasi Clustering Mahasiswa

Untuk menentukan cluster terlebih dahulu ditentukan jumlah cluster data, pada penelitian ini menggunakan 2 cluster yaitu mahasiswa berprestasi dan cukup berprestasi sedangkan penentuan cluster dapat dilihat pada tabel 2. berikut ini :

Tabel 2. Centroid 1 dan Centroid 2

<i>Centroid 1</i>				<i>Centroid 2</i>			
<i>IPK</i>	<i>Non-Akademik</i>	<i>Ikut Organisasi</i>	<i>Nilai Kemampuan Bahasa Asing</i>	<i>IPK</i>	<i>Non-Akademik</i>	<i>Ikut Organisasi</i>	<i>Nilai Kemampuan Bahasa Asing</i>
2.84	4	1	74	3.45	2	1	79

Setelah ditentukan centroid dengan tabel 2. diatas kemudian dihitung cluster 1 dan cluster 2 berdasarkan jarak terkecil data berikut perhitungan jarak antar data dapat dilihat pada tabel 3. Berikut :

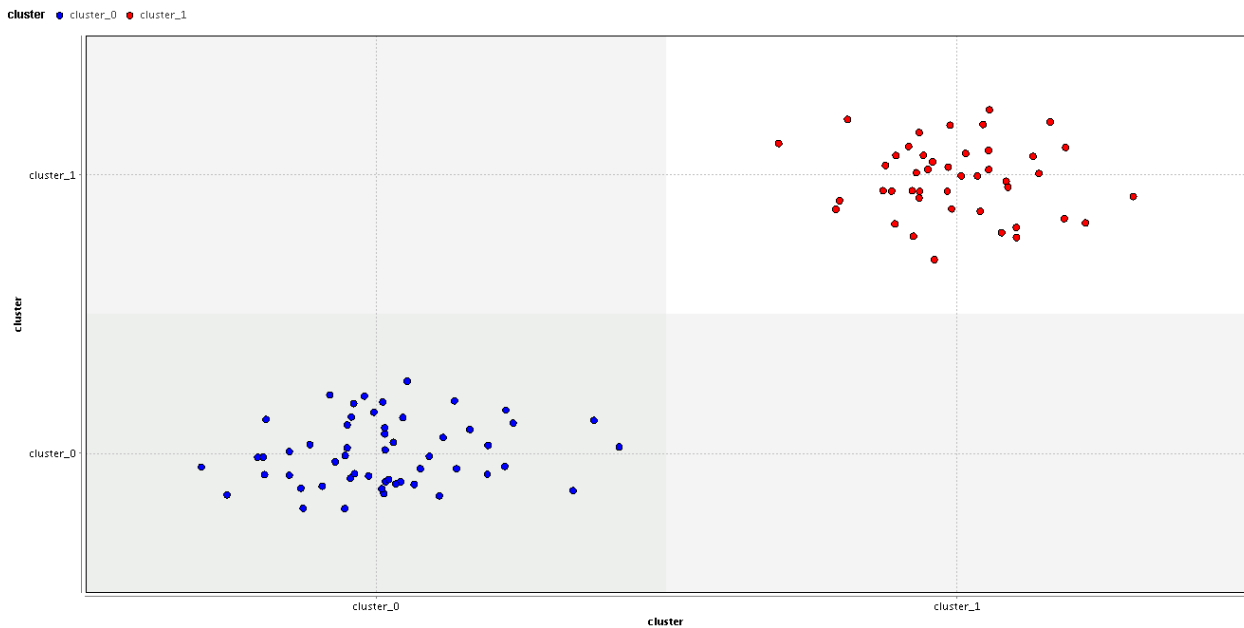
Tabel 3. Hasil perhitungan dan Hasil Cluster

<i>IPK</i>	<i>Non Akademik</i>	<i>Ikut Organisai</i>	<i>Kemampuan Bahasa</i>	<i>dc1</i>	<i>dc2</i>	<i>Cluster</i>
------------	---------------------	-----------------------	-------------------------	------------	------------	----------------

2.84	4	1	74	0.00	5.42	1
3.37	3	0	84	10.11	5.20	2
2.63	1	0	79	5.92	1.63	2
3	3	1	82	8.06	3.19	2
3.26	3	1	81	7.08	2.24	2
2.29	4	0	76	2.30	3.92	1
3.84	2	1	79	5.48	0.39	2
2.92	3	0	78	4.24	1.81	2
2.68	3	0	77	3.32	2.57	2
3.66	4	1	85	11.03	6.33	2
3.34	3	1	87	13.05	8.06	2
3.34	2	1	89	15.14	10.00	2
3.26	1	1	91	17.27	12.04	2
3.29	3	0	93	19.06	14.07	2
3.47	4	1	95	21.01	16.12	2
3.34	3	1	97	23.03	18.03	2
2.66	2	1	80	6.33	1.27	2
2.13	4	0	76	2.35	3.97	1
2.58	3	0	75	1.75	4.33	1
3.13	3	1	83	9.06	4.14	2
3.47	3	0	84	10.12	5.20	2
2.79	2	0	76	3.00	3.23	1
3.39	3	1	86	12.05	7.07	2
3.39	2	1	87	13.16	8.00	2
3.29	2	1	88	14.15	9.00	2
3.45	2	1	89	15.15	10.00	2
3.61	2	1	90	16.14	11.00	2
3.24	3	1	91	17.03	12.04	2
2.84	3	1	92	18.03	13.05	2
3.63	3	1	93	19.04	14.04	2
2.97	2	1	94	20.10	15.01	2
3.13	2	1	95	21.10	16.00	2
3.58	3	1	96	22.04	17.03	2
3.47	3	1	97	23.03	18.03	2
3.26	3	1	91	17.03	12.04	2
3.37	4	1	89	15.01	10.20	2
3.26	3	1	87	13.05	8.06	2
3.13	3	1	85	11.05	6.09	2
3.29	3	1	83	9.07	4.13	2
3.37	3	1	81	7.09	2.24	2
3.29	3	1	79	5.12	1.01	2
3.42	3	1	77	3.22	2.24	2
3.34	3	1	75	1.50	4.12	1
3.42	3	1	73	1.53	6.08	1
3.07	3	0	78	4.25	1.77	2
2.89	3	0	69	5.20	10.12	1
2.73	3	1	69	5.10	10.08	1
2.95	3	2	76	2.45	3.35	1
3.25	3	2	78	4.26	1.74	2
2.37	0	0	75	4.27	4.71	1
3.09	2	1	78	4.48	1.06	2
2.73	2	0	74	2.24	5.15	1

2.79	2	1	75	2.24	4.05	1
3.32	2	1	84	10.21	5.00	2
3.36	3	1	82	8.08	3.16	2
3.11	2	1	79	5.39	0.34	2
2.93	2	1	79	5.39	0.52	2
3.02	2	1	80	6.33	1.09	2
3.27	2	1	81	7.29	2.01	2
3.32	2	1	82	8.26	3.00	2
3.32	2	1	83	9.23	4.00	2
3.24	2	1	84	10.21	5.00	2
2.93	2	0	68	6.40	11.06	1
2.95	2	1	87	13.15	8.02	2
2.57	2	0	67	7.35	12.07	1
3	2	1	97	23.09	18.01	2
3.09	2	1	80	6.33	1.06	2
3.11	2	1	82	8.25	3.02	2
2.93	2	1	84	10.20	5.03	2
2.95	2	2	86	12.21	7.09	2
2.93	2	1	88	14.14	9.02	2
3.02	2	1	90	16.13	11.01	2
3.16	2	1	92	18.11	13.00	2
3.14	2	1	90	16.13	11.00	2
2.64	2	0	81	7.35	2.38	2
3.09	2	1	98	24.08	19.00	2
3.25	2	1	83	9.23	4.00	2
1.24	2	0	68	6.60	11.26	1
3.36	2	1	80	6.35	1.00	2
2	2	0	77	3.83	2.67	2
2.64	2	1	75	2.24	4.08	1
3	2	1	83	9.22	4.03	2
2.82	2	0	71	3.74	8.09	1
2.66	2	1	78	4.48	1.27	2
2.66	2	0	76	3.01	3.26	1
3.25	2	0	80	6.42	1.43	2
3.11	2	0	79	5.48	1.06	2
1.68	2	0	78	4.73	2.27	2
3.32	2	0	84	10.26	5.10	2
2.74	2	0	69	5.48	10.07	1
3.08	2	0	89	15.17	10.06	2
3.24	2	1	82	8.26	3.01	2
2.16	2	0	78	4.63	1.91	2
3.45	2	1	79	5.42	0.00	2
2.68	2	0	78	4.59	1.61	2
3.5	2	1	75	2.33	4.00	1
3.34	2	1	78	4.50	1.01	2
3.89	4	1	91	17.03	12.17	2

Pada tabel 3. Diatas dilakukan proses iterasi data sampai tidak ada perubahan lagi lakukan perulangan hingga nilai centroid yang dihasilkan tetap tidak berubah dan anggota cluster tidak berpindah ke cluster lain, setelah hasil cluster tetap itulah hasil cluster yang terbaik untuk memudahkan pengklasteran data dapat dilihat pada gambar 3. berikut :



Gambar 3. Hasil Cluster dalam Grafik

Pada gambar 3. diatas dapat dikelompokkan menjadi mahasiswa berprestasi dan tidak berprestasi mahasiswa beprestasi dalam cluster 1 dan mahasiswa tidak berprestasi dalam cluster 0

3.3. Pengujian Dengan *Davies Bouldin Index*

Untuk menguji apakah cluster yang telah dihasilkan baik atau tidak diperlukan pengujian yaitu dengan menggunakan metrik evaluasi *Davies Bouldin Index* yang digunakan untuk mengukur kualitas clustering nilai DBI yang lebih rendah menunjukkan clustering yang lebih baik kualitas clustering pada penelitian ini dengan hasil uji coba DBI sebesar 0.603 dan dianggap Cukup Baik.

4. KESIMPULAN

Identifikasi kelompok mahasiswa Melalui k-means clustering, mahasiswa dapat dikelompokkan berdasarkan Prestasi akademik dan non akademik Setiap kelompok mewakili tingkat prestasi yang berbeda. Penentuan faktor prestasi Melalui analisis hasil clustering, dan identifikasi faktor-faktor yang berkontribusi terhadap prestasi akademik mahasiswa, dengan memahami kelompok-kelompok mahasiswa yang berbeda, institusi pendidikan dapat mengambil keputusan yang lebih tepat dan spesifik untuk meningkatkan prestasi akademik mahasiswa. Misalnya, melalui program pembinaan tambahan untuk kelompok mahasiswa dengan nilai rendah atau memberikan kesempatan khusus bagi kelompok mahasiswa berprestas. Dalam kesimpulannya, clustering menggunakan metode k-means dapat membantu institusi pendidikan dalam mengidentifikasi dan memahami kelompok-kelompok mahasiswa berprestasi. Dengan pemahaman ini, institusi dapat mengambil langkah-langkah yang sesuai untuk meningkatkan prestasi akademik mahasiswa dan memonitor perubahan dari waktu ke waktu.

REFERENCES

- [1] P. Arora, Deepali, and S. Varshney, "Analysis of K-Means and K-Medoids Algorithm for Big Data," in *Physics Procedia*, Elsevier B.V., 2016, pp. 507–512. doi: 10.1016/j.procs.2016.02.095.
- [2] K. M. Kumar and A. R. M. Reddy, "An efficient k-means clustering filtering algorithm using density based initial cluster centers," *Inf Sci (N Y)*, vol. 418–419, pp. 286–301, Dec. 2017, doi: 10.1016/j.ins.2017.07.036.
- [3] B. Harahap, "Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung)".

- [4] L. Hakim, “PERANCANGAN APLIKASI PENILAIAN MAHASISWA BERPRESTASI UNIVERSITAS XYZ MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING Penulis Korespondensi,” 2013. [Online]. Available: <http://www.jurnal.umk.ac.id/sitech>
- [5] M. M. Öztürk, U. Cavusoglu, and A. Zengin, “A novel defect prediction method for web pages using k-means++,” *Expert Syst Appl*, vol. 42, no. 19, pp. 6496–6506, May 2015, doi: 10.1016/j.eswa.2015.03.013.
- [6] L. Keterampilan *et al.*, “CITRA ABDIMAS : Jurnal Pengabdian Kepada Masyarakat A B S T R A K Sejarah artikel,” vol. 2, no. 2, pp. 100–105, 2023, [Online]. Available: <https://publisher.yccm.or.id/index.php/cabp100Journalhomepage:https://publisher.yccm.or.id/index.php/cab>
- [7] Z. Kakushadze and W. Yu, “*K-means and cluster models for cancer signatures,” *Biomol Detect Quantif*, vol. 13, pp. 7–31, Sep. 2017, doi: 10.1016/j.bdq.2017.07.001.
- [8] M. Mughnyanti, S. Efendi, and M. Zarlis, “Analysis of determining centroid clustering x-means algorithm with davies-bouldin index evaluation,” in *IOP Conference Series: Materials Science and Engineering*, Institute of Physics Publishing, Jan. 2020. doi: 10.1088/1757-899X/725/1/012128.
- [9] M. Faisal, E. M. Zamzami, and Sutarman, “Comparative Analysis of Inter-Centroid K-Means Performance using Euclidean Distance, Canberra Distance and Manhattan Distance,” in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Jul. 2020. doi: 10.1088/1742-6596/1566/1/012112.
- [10] Y. Sopyan, A. D. Lesmana, and C. Juliane, “Analisis Algoritma K-Means dan Davies Bouldin Index dalam Mencari Cluster Terbaik Kasus Perceraian di Kabupaten Kuningan,” *Building of Informatics, Technology and Science (BITS)*, vol. 4, no. 3, Dec. 2022, doi: 10.47065/bits.v4i3.2697.
- [11] S. Paembonan, H. Abduh, and K. Kunci, “Penerapan Metode Silhouette Coeficient Untuk Evaluasi Clustering Obat Clustering; K-means; Silhouette coeficient,” 2021. [Online]. Available: <https://ojs.unanda.ac.id/index.php/jiit/index>