

ARTICLE

Implementasi Algoritma K-Means Clustering untuk Strategi Promosi Kampus IBISA

Implementation of K-Means Clustering Algorithm for IBISA Campus Promotion Strategy

Ummu Wachidatul Latifah,* Saiful Bahri, dan Marsita Satriandhini

Institut Teknologi Bisnis dan Kesehatan Bhakti Putra Bangsa Indonesia

*Penulis Korespondensi: ummuwl2@gmail.com

(Disubmit 24-04-18; Diterima 24-05-02; Dipublikasikan online pada 24-09-05)

Abstrak

Penerimaan Mahasiswa Baru (PMB) adalah kegiatan yang wajib dilaksanakan di setiap universitas, seperti halnya pada IBISA. Namun, di Jawa Tengah terdapat beberapa kampus yang melakukan PMB. Oleh karena itu, perlu adanya strategi promosi untuk mendapatkan jumlah mahasiswa baru yang lebih banyak. Analisa strategi promosi IBISA yang sesuai dan akurat dalam mendapatkan mahasiswa baru sangat diperlukan. Teknik promosi yang baik dapat ditentukan menggunakan algoritma clustering. Algoritma clustering yang dapat diimplementasikan, yaitu algoritma K-Means. Metode penelitian yang digunakan, yaitu metode kuantitatif. Objek penelitian berupa data mahasiswa baru IBISA tahun 2022 dan 2023. Hasil dari penelitian ini, yaitu diperoleh 2 cluster. Pengujian dilakukan tiga kali dengan nilai centroid awal yang ditentukan secara acak. Hasil dari tiga kali pengujian memberikan hasil yang akurat dan stabil. Berdasarkan hal tersebut dapat dinyatakan bahwa bahwa strategi promosi kampus IBISA perlu ditingkatkan lagi, yaitu dengan melakukan promosi atau pengenalan kampus IBISA di luar Purworejo ataupun Luar Jawa.

Kata kunci: Strategi Promosi; Perguruan Tinggi; Clustering; K-Means

Abstract

New Student Admission (PMB) is an activity that must be carried out at every university, as is the case with IBISA. However, in Central Java there are few campuses that conduct PMB. Therefore, it is necessary to have a promotional strategy to get more new students. Analysis of IBISA promotion strategies that are appropriate and accurate in getting new students is needed. A good promotion technique can be determined using a clustering algorithm. The clustering algorithm that can be implemented is the K-Means algorithm. The research method used is quantitative method. The object of research is IBISA new student data in 2022 and 2023. The results of this study, namely obtained 2 clusters. Testing was done three times with the initial centroid value determined randomly. The results of the three tests gave accurate and stable results. Based on this, it can be stated that the IBISA campus promotion strategy needs to be improved again, namely by promoting or introducing the IBISA campus outside Purworejo or outside Java.

KeyWords: Promotion Strategy, Higher Education, Clustering, K-Means

1. Pendahuluan

Pendidikan adalah faktor penting sebagai upaya membangun bangsa dan memajukan suatu negara. Melalui pendidikan seseorang menjadi cerdas, memiliki kapasitas dan mentalitas unggul untuk hidup, dan mulai dapat mengurangi kebutuhan hidup ke tingkat yang lebih tinggi. Berdasarkan Pembukaan UUD

This is an Open Access article - copyright on authors, distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International License (CC BY SA) (<http://creativecommons.org/licenses/by-sa/4.0/>)

How to Cite: U.W. Latifah *et al.*, "Implementasi Algoritma K-Means Clustering untuk Strategi Promosi Kampus IBISA", *JIKO (JURNAL INFORMATIKA DAN KOMPUTER)*, Volume: 8, No.2, Pages 292–300, September 2024, doi: 10.26798/jiko.v8i2.1307.

1945 tujuan pendidikan nasional, yaitu untuk meningkatkan kehidupan bangsa. Hal tersebut mencakup kecerdasan intelektual dan kecerdasan memahami makna lebih luas[1]. Untuk mencapai tujuan pendidikan, maka diperlukan pembentukan pendidikan berjenjang. Pendidikan berjenjang dimulai dari sekolah dasar, sekolah menengah pertama, sekolah menengah atas dan pendidikan tinggi.

Penyelenggaraan perguruan tinggi merupakan salah satu upaya untuk meningkatkan kemampuan akademik dan keahlian profesi sehingga Para siswa dari kelas bawah yang melanjutkan pendidikan di tingkat atas memiliki pemahaman yang lebih dalam dan lebih percaya diri[2]. Perguruan tinggi di Indonesia terdiri dari perguruan tinggi negeri (PTN) ataupun perguruan tinggi swasta (PTS). PTN ataupun PTS diharapkan mampu memberikan pendidikan yang sesuai dengan kemajuan teknologi dan informasi. Salah satu PTS yang dapat memberikan pendidikan yang dibutuhkan, yaitu IBISA.

IBISA merupakan singkatan dari Institut Teknologi Bisnis dan Kesehatan Bhakti Putra Bangsa Indonesia yang terletak di kecamatan Banyuwirip kabupaten Purworejo Jawa Tengah. IBISA merupakan salah satu kampus yang menerima mahasiswa baru setiap tahunnya.

Penerimaan Mahasiswa Baru (PMB) merupakan agenda yang dilakukan di beberapa universitas, termasuk di kampus IBISA. Namun, di Jawa Tengah terdapat beberapa kampus yang melakukan PMB. Jawa. PMB yang selalu diselenggarakan tiap tahun mengumpulkan informasi calon Mahasiswa baru dari berbagai daerah. Berdasarkan informasi yang dieproleh dapat diambil kesimpulan dengan cara mengolah data atau informasi supaya mendapatkan informasi – informasi baru. Informasi yang diperoleh dapat dipakai sebagai penunjang keputusan bagi universitas[3]. Oleh karena itu, perlu adanya strategi promosi untuk menentukan berapa banyak siswa baru yang diterima. Memeriksa kampanye iklan IBISA dengan cara yang benar dan jujur dalam menerima siswa baru diperlukan. Lokasi sekolah yang dikunjungi belum ditentukan secara optimal dan tidak melakukan pengujian tanah dengan potensi tinggi. Kurangnya anggaran dan waktu adalah ketika proses strategi promosi tidak dilakukan dengan baik dan efektif[4].

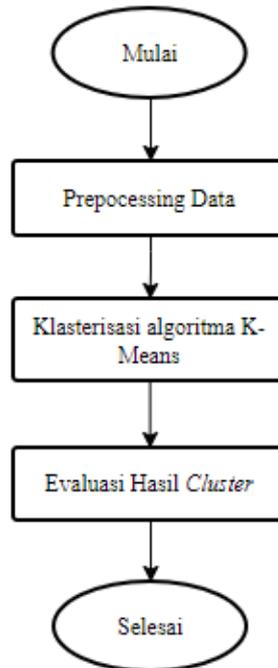
Berdasarkan permasalahan yang ada, diperlukan suatu metode analisis *cluster* yang dapat memaksimalkan teknik promosi. Salah satu algoritma *clustering* yang dapat diterapkan, yaitu algoritma K-Means. Pemilihan strategi promosi dilakukan dengan menggunakan K-means *Clustering*, yaitu, data-data yang diperoleh Berdasarkan karakteristik ini, data dikelompokkan ke dalam banyak kelompok, dengan data yang memiliki karakteristik yang sama dimasukkan ke dalam satu kelompok dan data yang memiliki karakteristik yang berbeda dimasukkan ke dalam kelompok lain.

Konsep dasar dari algoritma K-Means dimulai dari menentukan jumlah grup (k), lalu pilih pusat cluster (centroid) secara acak. Kedua, menghitung jarak dari setiap data ke pusat cluster. Ketiga, data dikelompokkan ke dalam setiap cluster dengan jarak terdekat. Keempat, menghitung pusat grup baru kemudian mengulangi langkah hingga pusat cluster tidak berubah atau mencapai maksimum iterasi. Dalam proses clustering dapat dimulai dengan mengidentifikasi data yang dikelompokkan, menggunakan rumus Euclidean Distance[5]. Kelebihan K-Means, yaitu mudah diterapkan dan waktu proses relatif cepat[6]. Namun, algoritma K-Means memiliki kelemahan atau permasalahan. Permasalahan yang ada, yaitu penentuan jumlah cluster yang kurang ideal. Hal tersebut dapat menghasilkan centroid akhir yang tidak benar-benar menjadi pusat cluster yang sesungguhnya. Metode evaluasi cluster dapat menyelesaikan masalah tersebut salah satunya, yaitu metode Elbow[7]. Oleh karena itu, pada penelitian ini akan digunakan metode Elbow untuk menentukan jumlah cluster optimal.

Penelitian lain yang terkait promosi dengan metode K-Means, yaitu penelitian yang dilakukan oleh[2]. Hasil analisis dengan algoritma K-Means menunjukkan bahwa algoritma K-Means dapat digunakan untuk memilih program studi. Metode posisi kompetitif adalah strategi yang efektif dan media promosi harus menunjukkan keunggulan program studi. Penelitian lain dilakukan oleh[8] memperoleh hasil strategi untuk menentukan kontribusi dari kampus, seperti pelatihan IT, simulasi try out, dan memberikan sponsor edukasi IT sesuai kebutuhan sekolah. Penelitian yang dilakukan[9] Hasil dari penelitian ini digunakan untuk menentukan strategi periklanan berdasarkan berbagai cluster media promosi. Penelitian yang dilakukan terkait promosi kampus IBISA menggunakan algoritma clustering K-Means dalam menentukan strategi promosi kampus IBISA untuk memperoleh mahasiswa dalam proses PMB.

2. Metode

Metode penelitian yang digunakan pada penelitian ini, yaitu metode kuantitatif. Metode kuantitatif merupakan metode yang digunakan pada suatu penelitian dimana data yang akan dianalisis berupa angka (numerik). Penelitian ini mengambil objek dari data mahasiswa baru IBISA tahun 2022 dan 2023. Variable penelitian yang digunakan berupa asal siswa atau alamat rumah siswa. Langkah-langkah penelitian ini dapat dilihat pada Gambar 1 berikut ini:



Gambar 1. Flowchart Tahapan Penelitian

2.1 Metode Elbow

Metode Elbow adalah teknik untuk menentukan ukuran kluster yang optimal dengan membandingkan ukuran relatif dari berbagai ukuran kluster dipresentasikan dengan bentuk siku di suatu titik. Metode Elbow merupakan salah satu metode untuk memperbaiki kelemahan algoritma K-Means *clustering* dalam penentuan jumlah cluster ideal. Metode ini digunakan untuk menguji konsistensi jumlah cluster terbaik dengan cara membandingkan selisih *Sum of Square Error* (SSE) masing-masing *cluster*, selisih paling ekstrim membentuk sudut siku-siku menunjukkan jumlah cluster terbaik[10]. Berikut ini merupakan persamaan metode Elbow[9]:

$$SSE = \sum_{i=1}^k \sum_{x_j \in C_i} |X_i - \mu_i|^2 \quad (1)$$

dengan,

k= jumlah cluster

X_i = data ke-i

μ_i = titik pusat cluster

2.2 Algoritma K-Means Clustering

Langkah-langkah dalam algoritma K-Means clustering, sebagai berikut[5]:

1. Menentukan jumlah cluster K dan jumlah iterasi maksimum.

2. Menentukan titik pusat (centroid). Nilai centroid awal ditentukan secara acak, untuk iterasi selanjutnya menggunakan persamaan berikut ini:

$$C_i = \frac{1}{M} \sum_{j=1}^M x_j \tag{2}$$

dimana:

C_i = rata-rata cluster ke i

M= banyak data yang menjadi anggota cluster ke i

i,j= indeks cluster

x_j = nilai data ke j yang ada di cluster tersebut

3. Menghitung jarak antara titik pusat dan titik setiap objek. Menghitung jarak menggunakan persamaan Euclidean Distance, yaitu:

$$d = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{3}$$

dimana:

d = rata-rata cluster ke i

$(x_i - y_i)$ = Koordinat objek

$(x_j - y_j)$ = Koordinat centroid

4. Realokasi data ke masing-masing cluster dilakukan berdasarkan perbandingan jarak antara data dengan centroid masing-masing cluster. Penentuan anggota cluster dilakukan dengan mempertimbangkan jarak minimum objek. Nilai yang diperoleh pada keanggotaan jarak adalah 0 atau 1. Nilai 1 untuk data yang dialokasikan pada cluster tersebut, sedangkan nilai 0 untuk data yang dialokasikan pada cluster lain.

$$a_{ij} \begin{cases} 1 & d = \text{Min} \{D(x_i, c_i)\} \\ 0 & \text{lainnya} \end{cases} \tag{4}$$

5. Hitung ulang posisi centroid. a_{ij} merupakan nilai keanggotaan x_i ke titik pusat cluster c_1 , d adalah jarak terdekat dari titik x_i ke cluster K setelah dibandingkan dan c_1 merupakan pusat cluster ke 1. Fungsi objektif K-Means Komposisi grup ditentukan oleh peringkat dan nilai anggota datanya. Fungsi objektif dapat ditentukan menggunakan persamaan[10]:

$$J = \sum_{i=1}^n \sum_{l=1}^k a_{il} cD(x_i c_i)^2 \tag{5}$$

n adalah jumlah data, k jumlah cluster, a_{il} nilai keanggotaan dari data x_i ke cluster c_l yang diikuti a, a mempunyai dua kemungkinan, yaitu bernilai 1 atau 0. Jika a bernilai 1 maka data merupakan anggota kelompok tersebut dan apabila a bernilai 0 maka data dialokasikan ke cluster lain.

6. Apabila terjadi perubahan centroid atau jumlah iterasi kurang dari jumlah iterasi maksimum maka iterasi dilanjutkan, jika tidak maka iterasi berhenti dan diperoleh hasil cluster.

3. Hasil dan Pembahasan

Pada sub bagian ini, akan diuraikan hasil pengujian algoritma K-Means yang diimplementasikan pada data pendaftar mahasiswa baru IBISA tahun 2022 dan 2023. Variable yang digunakan adalah asal kota siswa. Data tersaji pada Tabel[1].

Tabel 1. Data Mahasiswa Baru IBISA 2022

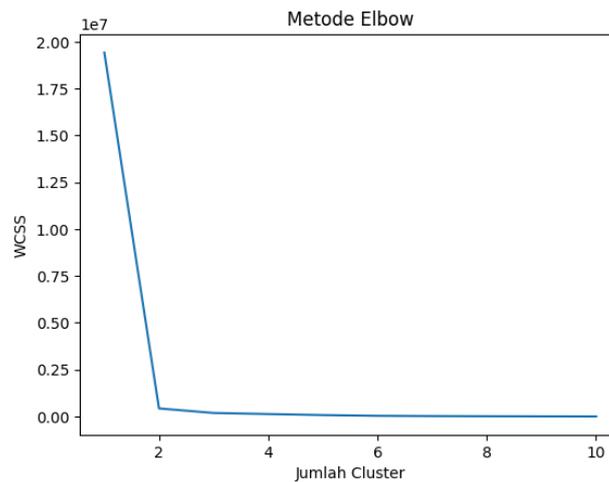
Nama	Kota/Kab	Jarak (Km)
Budiman Sujatmiko	Purworejo	4,2
Mutafaqih Fidin	Purworejo	12
Steavani Vernando	Banjarnegara	69
...
Wiji Sugiyati	Purworejo	12

Tabel 2. Data Mahasiswa Baru IBISA 2023

Nama	Kota/Kab	Jarak (Km)
Dian Puspitasari	Purworejo	16,00
Nanda Eka Pratiwi	Purworejo	10,00
Annisak Asmaul Husna	Purworejo	8,00
...
Fadhillah Indah S	Purworejo	5,00

3.1 Metode Elbow

Langkah selanjutnya, yaitu menentukan jumlah cluster terbaik. Pada penelitian ini, penentuan cluster terbaik menggunakan metode Elbow. Penentuan cluster terbaik menggunakan persamaan metode Elbow yang dihitung menggunakan Python. Hasil penentuan cluster terbaik menggunakan metode Elbow dapat dilihat pada Gambar 2.

**Gambar 2.** Jumlah Cluster Terbaik Menggunakan Metode Elbow

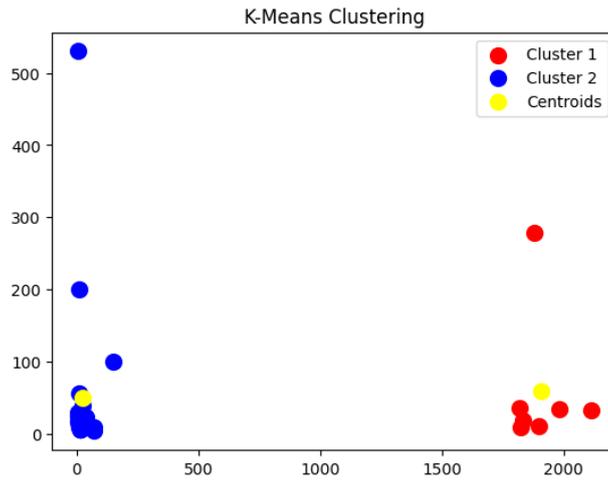
Berdasarkan metode Elbow terlihat bahwa nilai 2 membentuk titik siku sehingga nilai 2 merupakan jumlah cluster optimum atau terbaik [11].

3.2 Algoritma K-Means Clustering

Langkah awal analisis K-Means clustering, yaitu menentukan jumlah cluster terbaik menggunakan metode Elbow. Berdasarkan metode Elbow diperoleh jumlah cluster terbaik adalah 2 cluster. Berikutnya, yaitu melakukan analisis cluster menggunakan metode K-Means dengan jumlah cluster 2. Hasil dari K-Means sangat berpengaruh terhadap centroid yang dipilih, maka untuk memperoleh hasil yang akurat analisa akan dilakukan sebanyak 3 kali pengujian dengan menggunakan nilai centroid yang berbeda.

3.2.1 K-Means Clustering Pengujian Pertama

Hasil K-Means clustering pada pengujian pertama ditampilkan dalam gambar di bawah ini.



Gambar 3. Plot Hasil K-Means Clustering Pengujian Pertama

Pada Gambar 3 terlihat bahwa cluster 1 merupakan cluster rendah dengan jumlah sebanyak 7 anggota, dan cluster 2 adalah cluster tinggi dengan jumlah sebanyak 23 anggota. Nilai centroid hasil pengujian pertama disajikan pada Tabel 3.

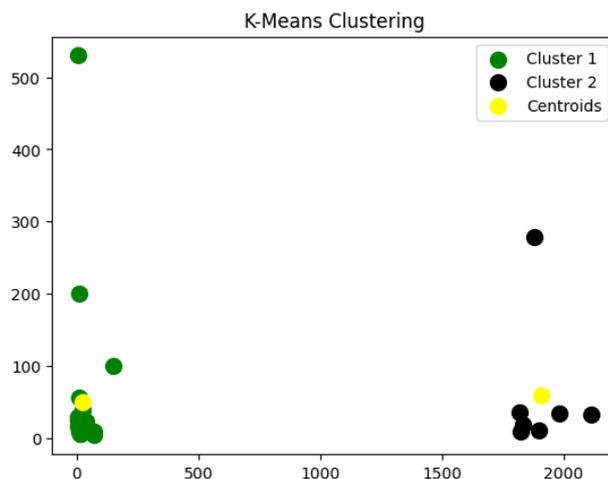
Tabel 3. Nilai Centroid Pengujian Pertama

Cluster	X_1	X_2
1	1906,42	59,57
2	25,20	50,05

Berdasarkan Tabel 3 dapat terlihat bahwa cluster 1 merupakan cluster yang memiliki sedikit anggota, dengan nilai centroid 1906.42 dan 59.57. Hal tersebut menyatakan bahwa rata-rata jarak mahasiswa yang termasuk dalam cluster 1 merupakan mahasiswa yang berasal dari luar Purworejo ataupun luar Pulau Jawa. Sedangkan cluster 2 dengan nilai centroid 25.20 dan 50.05 merupakan cluster yang memiliki banyak anggota. Hal tersebut menyatakan bahwa rata-rata jarak mahasiswa yang termasuk dalam cluster 2 merupakan mahasiswa yang berasal dari Purworejo ataupun dalam Pulau Jawa.

3.2.2 K-Means Clustering Pengujian Kedua

Hasil K-Means clustering pada pengujian kedua ditampilkan dalam gambar di bawah ini.



Gambar 4. Plot Hasil K-Means Clustering Pengujian Kedua

Pada Gambar 4 terlihat bahwa cluster 1 merupakan cluster tinggi dengan jumlah sebanyak 23 anggota, dan cluster 2 adalah cluster rendah dengan jumlah sebanyak 7 anggota. Nilai centroid hasil pengujian kedua disajikan pada Tabel 4.

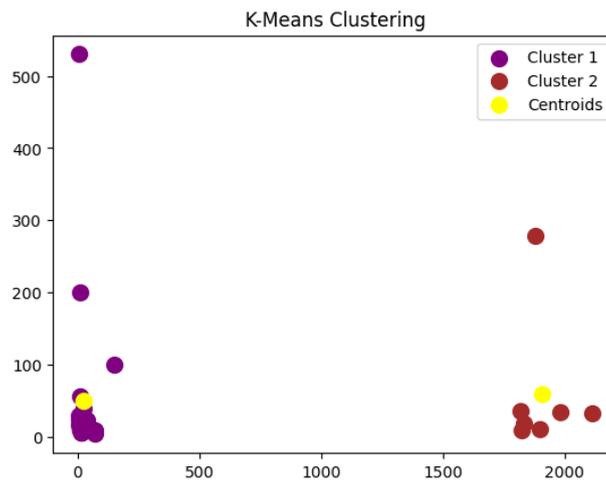
Tabel 4. Nilai Centroid Pengujian Kedua

Cluster	X_1	X_2
1	25,20	50,05
2	1906,42	59,57

Berdasarkan Tabel 4 dapat terlihat bahwa cluster 1 merupakan cluster yang memiliki banyak anggota, dengan nilai centroid 25.20 dan 50.05. Hal tersebut menyatakan bahwa rata-rata jarak mahasiswa yang termasuk dalam cluster 1 merupakan mahasiswa yang berasal dari Purworejo ataupun Pulau Jawa. Sedangkan cluster 2 dengan nilai centroid 1906.42 dan 59.57 merupakan cluster yang memiliki sedikit anggota. Hal tersebut menyatakan bahwa rata-rata jarak mahasiswa yang termasuk dalam cluster 2 merupakan mahasiswa yang berasal dari luar Purworejo ataupun luar Pulau Jawa.

3.2.3 K-Means Clustering Pengujian Ketiga

Hasil K-Means clustering pada pengujian ketiga ditampilkan dalam gambar di bawah ini.



Gambar 5. Plot Hasil K-Means Clustering Pengujian Ketiga

Pada Gambar 5 terlihat bahwa cluster 1 merupakan cluster tinggi dengan jumlah sebanyak 23 anggota, dan cluster 2 adalah cluster rendah dengan jumlah sebanyak 7 anggota. Nilai centroid hasil pengujian ketiga disajikan pada Tabel 5.

Tabel 5. Nilai Centroid Pengujian Ketiga

Cluster	X_1	X_2
1	25,20	50,05
2	1906,42	59,57

Berdasarkan Tabel 5 dapat terlihat bahwa cluster 1 merupakan cluster yang memiliki banyak anggota, dengan nilai centroid 25.20 dan 50.05. Hal tersebut menyatakan bahwa rata-rata jarak mahasiswa yang termasuk dalam cluster 1 merupakan mahasiswa yang berasal dari Purworejo ataupun Pulau Jawa. Sedangkan cluster 2 dengan nilai centroid 1906.42 dan 59.57 merupakan cluster yang memiliki sedikit anggota. Hal tersebut menyatakan bahwa rata-rata jarak mahasiswa yang termasuk dalam cluster 2 merupakan mahasiswa yang berasal dari luar Purworejo ataupun luar Pulau Jawa.

4. Simpulan

Hasil dari penelitian ini diperoleh jumlah cluster terbaik yang dari metode Elbow adalah 2. Pengujian yang telah dilakukan tiga kali dengan nilai centroid yang berbeda, untuk memperoleh hasil akurat. Berdasarkan hasil dan pembahasan menyatakan bahwa hasil yang diperoleh tetap akurat dan memiliki nilai centroids yang sama dalam setiap pengujian. Hasil pengujian menyatakan adanya perbedaan pada penomoran cluster, namun hal tersebut tidak mempengaruhi hasil pengelompokan dan nilai centroids yang dihasilkan. Pada pengujian pertama diperoleh bahwa cluster 1 merupakan cluster yang memiliki sedikit anggota, dengan nilai centroid 1906.42 dan 59.57. Hal tersebut menyatakan bahwa rata-rata jarak mahasiswa yang termasuk dalam cluster 1 merupakan mahasiswa yang berasal dari luar Purworejo ataupun luar Pulau Jawa. Sedangkan cluster 2 dengan nilai centroid 25.20 dan 50.05 merupakan cluster yang memiliki banyak anggota. Hal tersebut menyatakan bahwa rata-rata jarak mahasiswa yang termasuk dalam cluster 2 merupakan mahasiswa yang berasal dari Purworejo ataupun dalam Pulau Jawa. Pada pengujian kedua dan ketiga diperoleh hasil yang sama, yaitu cluster 1 merupakan cluster yang memiliki banyak anggota, dengan nilai centroid 25.20 dan 50.05. Hal tersebut menyatakan bahwa rata-rata jarak mahasiswa yang termasuk dalam cluster 1 merupakan mahasiswa yang berasal dari Purworejo ataupun Pulau Jawa. Sedangkan cluster 2 dengan nilai centroid 1906.42 dan 59.57 merupakan cluster yang memiliki sedikit anggota. Hal tersebut menyatakan bahwa rata-rata jarak mahasiswa yang termasuk dalam cluster 2 merupakan mahasiswa yang berasal dari luar Purworejo ataupun luar Pulau Jawa. Berdasarkan hal tersebut dapat disimpulkan bahwa strategi promosi kampus IBISA perlu ditingkatkan lagi sehingga dapat menyebar ke luar Purworejo dan luar Jawa.

5. Saran

Saran untuk penelitian selanjutnya, yaitu meminimalisir kesalahan Algoritma K-Means Clustering dalam penentuan centroids awal dapat menggunakan Genetic Algorithm.

Pustaka

- [1] B. Yusuf and A. Aziizu, "46 tujuan besar pendidikan adalah tindakan," [Online]. Available: <http://akhmadsudrajat>.
- [2] B. Nurhakim and K. Anam, "Penerapan strategi promosi kampus menggunakan k-means di stmik ikmi cirebon," *Jurnal ICT: Information Communication & Technology*, vol. 21, pp. 279–285, 2022.
- [3] A. L. Maukar, F. Marisa, A. A. Widodo, N. Kamilaningtyas, D. Novian, and D. Nugraha, "Analisis data penerimaan mahasiswa baru berbasis k-means," *Jurnal Informatika dan Komputer*, vol. 6, no. 2, pp. 142–147, 2022.
- [4] A. Bellanov and L. Nurhayati, "K-means clustering analysis untuk menentukan strategi promosi kampus," 2023.
- [5] U. S. S. W. Latifah, "K-means and fuzzy c-means algorithm comparison on regency/city grouping in central java province," *Desimal: Jurnal Matematika*, vol. 5, no. 2, pp. 155–168, 2022.
- [6] M. A. Soeleman and F. Ilmu Komputer, "Penentuan centroid awal pada algoritma k-means dengan dynamic artificial chromosomes genetic algorithm untuk tuberculosis dataset pre-centroid determination in k-means algorithm using dynamic artificial chromosomes genetic algorithm for tuberculosis dataset."
- [7] M. Orisa, "Optimasi cluster pada algoritma k-means," vol. 13, 2022.
- [8] A. Ramadhani, R. Fauziah, and S. Royal, "Clustering tingkat promosi kampus dengan menggunakan algoritma k-means," 2021, [Online]. Available: <http://jurnal.goretanpena.com/index.php/JSSR>.
- [9] I. Mahmudi, A. D. Indriyanti, and I. Lazulfa, "Penerapan algoritma k-means clustering sebagai strategi promosi penerimaan mahasiswa baru pada universitas hasyim asy'ari jombang."

- [10] E. Umargono, J. E. Suseno, and V. Gunawan, "K-means clustering optimization using the elbow method and early centroid determination based on mean and median formula," 2020.
- [11] R. Nainggolan, R. Perangin-Angin, E. Simarmata, and A. F. Tarigan, "Improved the performance of the k-means cluster using the sum of squared error (sse) optimized by using the elbow method," in *Journal of Physics: Conference Series*. Institute of Physics Publishing, Dec 2019.