

KLASTERISASI MAHASISWA MAGETAN MENGGUNAKAN K-MEANS UNTUK OPTIMASI STRATEGI PROMOSI PERGURUAN TINGGI

Aqsa Prima Cahya¹⁾, Muhammad Asyraf²⁾, Yudhistira Nanda Kumala³⁾,
Eva Yulia Puspaningrum⁴⁾

E-mail : ¹⁾22081010059@student.upnjatim.ac.id , ²⁾22081010107@student.upnjatim.ac.id ,
³⁾22081010055@student.upnjatim.ac.id , ⁴⁾evayuliapuspaningrum.if@upnjatim.ac.id

^{1,2,3,4)}Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jawa Timur

(Naskah masuk: Desember 2024, diterima untuk diterbitkan: 30 April 2025)

Abstrak

Semakin bertambahnya tahun, data mahasiswa magetan akan terus bertambah hingga menghasilkan tumpukan data yang berlimpah. Perlu adanya pengolahan data sehingga tumpukan data tersebut dapat dimanfaatkan sebagai ladang informasi. Penelitian ini bertujuan untuk mengkluster data mahasiswa Kabupaten Magetan yang sedang berkuliah di Universitas yang ada di Surabaya melalui proses data mining dengan algoritma *K-Means* serta metode *Elbow* dan *Silhouette Coefficient* dalam pembentukan clusternya. Data atribut yang digunakan pada penelitian ialah nama, asal sekolah, dan juga universitas. Data bersumber dari mahasiswa sendiri melalui pengisian google form oleh Organisasi Ikatan Mahasiswa Magetan di Surabaya, dimana data yang digunakan merupakan data mahasiswa angkatan 2023 dan 2024 dengan total sampel data sebanyak 250 *items*. Setelah melakukan perhitungan dengan metode *Elbow* didapatkan jumlah *cluster* sebanyak 4. Kemudian dilakukan evaluasi menggunakan metode *Silhouette Coefficient* dan didapatkan rata-rata terdekat dari nilai 1 adalah *cluster* 2, dengan nilai 0,62. Karena kohesivitas yang lebih baik serta model yang lebih sederhana, hasil *cluster* yang paling optimal adalah sebanyak 2 *cluster* pada *epoch* ke-5 dengan *cluster* 1 sebanyak 65 *items*, dan *cluster* 2 sebanyak 160 *items*. Adanya penelitian ini diharapkan dapat membantu universitas yang ada di Surabaya untuk menunjang strategi promosi berdasarkan hasil *cluster* universitas yang banyak diminati dari masing-masing sekolah di Kabupaten Magetan.

Kata kunci: *data mining, K-means, clustering, silhouette coefficient, elbow.*

1. PENDAHULUAN

Perkembangan teknologi informasi yang pesat telah menghasilkan akumulasi data dalam jumlah besar di berbagai bidang, termasuk pendidikan tinggi. Data mahasiswa yang terus bertambah setiap tahun ajaran baru menghasilkan tumpukan informasi yang berpotensi memberikan wawasan berharga jika diolah dengan tepat. Salah satu tantangan yang dihadapi oleh perguruan tinggi adalah bagaimana memanfaatkan data tersebut untuk mendukung strategi promosi yang efektif dalam menarik calon mahasiswa baru. Kabupaten Magetan, sebagai salah satu wilayah di Jawa Timur, memiliki potensi besar dalam menyumbang mahasiswa ke berbagai perguruan tinggi di Surabaya. Namun, pola pemilihan universitas oleh siswa dari Magetan belum dipetakan secara komprehensif. Pemahaman terhadap preferensi ini dapat membantu perguruan tinggi di Surabaya untuk merancang strategi promosi yang lebih terarah dan efisien.

Penelitian tentang analisis data mahasiswa menggunakan teknik data mining sebagai bantuan dalam strategi promosi perguruan tinggi sudah pernah dilakukan sebelumnya, meskipun dengan data mahasiswa dan algoritma yang berbeda. Penelitian-penelitian

terdahulu ini dapat dijadikan bahan pertimbangan dan diharapkan mampu memberikan kontribusi dalam penelitian yang sedang dilakukan. Penelitian oleh Hariri & Ramadhani

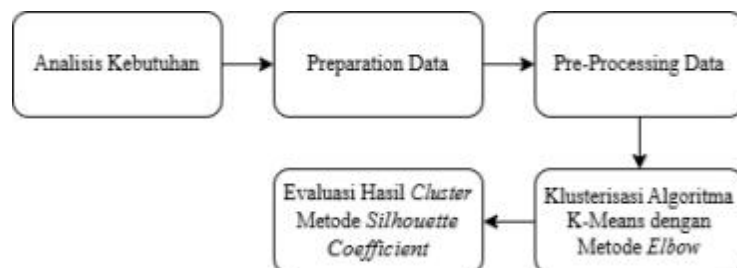
menggunakan metode *Association Rules* untuk menganalisis hubungan antara daerah asal dan minat mahasiswa terhadap program studi di Universitas Nusantara PGRI Kediri [1]. Hasil dari penelitian tersebut digunakan oleh pengambil keputusan untuk memberikan rekomendasi wilayah promosi yang lebih efektif. Penelitian oleh Alfina menggunakan metode *Hierarchical Clustering*, *K-means* dan Gabungan Keduanya pada data problem kerja praktek Jurusan Teknik Industri ITS untuk menjelaskan bagaimana cara mengelompokkan problem kerja praktek berdasarkan posting problem yang ada pada forum diskusi online SI-KP [2]. Penelitian oleh Lestari menggunakan algoritma *K-Means Clustering* pada data penerimaan mahasiswa baru STMIK Bina Bangsa Kendari dengan tujuan membantu bagian promosi STMIK Bina Bangsa Kendari dalam melakukan promosi dan mencari calon mahasiswa baru [3].

Dengan mempertimbangkan uraian di atas serta referensi dari penelitian-penelitian sebelumnya, maka teknik data mining yang akan digunakan pada penelitian ini adalah *clustering* dengan metode *K-means*. Hal ini dikarenakan *K-means* mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien. Namun, *K-means* mempunyai kelemahan yang diakibatkan oleh penentuan pusat awal *cluster*. Hasil *cluster* yang terbentuk dari metode *K-means* ini sangatlah tergantung pada inisiasi nilai pusat awal *cluster* yang diberikan [1]. Oleh karena itu akan digunakan juga metode *elbow* untuk menentukan jumlah *cluster* yang tepat melalui persentase hasil perbandingan antara jumlah *cluster* yang akan membentuk siku pada suatu titik.

Penelitian ini bertujuan untuk mengkluster data mahasiswa Kabupaten Magetan yang sedang berkuliah di Universitas yang ada di Surabaya melalui proses data mining dengan algoritma *K-Means* serta metode *Elbow* dan *Silhouette* dalam pembentukan klasternya. Data atribut yang digunakan pada penelitian ialah nama, asal sekolah, dan juga universitas. Data bersumber dari mahasiswa sendiri melalui pengisian google form oleh Ikatan Mahasiswa Magetan di Surabaya, dimana data yang digunakan merupakan data mahasiswa angkatan 2023 dan 2024 dengan total sampel data sebanyak 250 *items*. Adanya penelitian ini diharapkan dapat membantu perguruan tinggi yang ada di Surabaya untuk menunjang strategi promosi berdasarkan hasil *cluster* universitas yang banyak diminati dari masing-masing sekolah di Kabupaten Magetan. Hasil *clustering* ini diharapkan dapat memberikan gambaran yang jelas tentang kecenderungan pemilihan universitas oleh siswa dari sekolah-sekolah tertentu di Kabupaten Magetan.

2. METODOLOGI

Metode yang digunakan pada penelitian ini ialah metode kuantitatif. Metode kuantitatif sendiri ialah metode yang digunakan pada suatu penelitian dimana data yang akan digunakan berupa data yang berbentuk angka (numerik) [4]. Meskipun data awal yang digunakan akan berbentuk kualitatif, namun setelah tahap pre-processing akan berubah menjadi data numerik. Sedangkan untuk langkah-langkah pada penelitian ini dapat dilihat pada gambar 1 berikut :



Gambar 1. Alur Metodologi Penelitian

2.1 Analisis Kebutuhan

Tahapan pertama pada penelitian ialah analisis kebutuhan atau biasa juga disebut identifikasi masalah. Identifikasi masalah bermaksud supaya permasalahan dapat dirumuskan sesuai dengan apa yang akan dikaji, menentukan batasan masalah, juga menjelaskan manfaat dan tujuan penelitian [5].

2.2 Preparation Data

Pada tahapan ini akan dikumpulkamm data melalui metode kuisioner dimana data didapatkan melalui pengisian secara langsung bagi mahasiswa Magetan angkatan 2023 dna 2024 yang sedang berkuliah di Surabaya pada platform Google Form. Selain itu, akan dilakukan *Selection Data* yakni menentukan data yang akan diambil untuk menghindari kesamaan dan perulangan data [6].

2.3 Pre-Processing Data

Terdapat dua tahapan yakni *Data Cleaning* dan *Data Integration*. Data Cleaning ialah menghilangkan data yang rusak seperti *missing value*, *noise* ataupun mengangani data yang tidak konsisten. Data Integration digunakan untuk entitas yang bersifat unik [6]. Serta kemudian untuk lebih menyederhanakan data, perlu adanya transformasi data, yakni mengubah data yang berjenis inisialisasi nama dirubah menjadi singkatan [7].

2.4 Klusterisasi Algoritma K-Means

Data mining ialah proses ekstraksi analisis dari data kompleks untuk identifikasi pola, informasi, dan pengetahuan yang berharga dan tersembunyi [8]. Sedangkan clustering menurut Widodo ialah metode untuk mengelompokkan data berdasarkan kedekatan dengan pusat *cluster* [9]. Pada penelitian ini akan digunakan Klusterisasi dengan algoritma *K-Means*. Algoritma *K-Means* adalah salah satu teknik clustering non-hirarkis yang membagi data menjadi beberapa kelompok berdasarkan kesamaan karakteristik [10].

2.5 Evaluasi Hasil Kluster

Pengujian data dilakukan menggunakan metode *Elbow* dan *Silhouette Coefficient* untuk mengevaluasi kualitas *clustering* dengan menganalisis jumlah dan fitur turunan dari kumpulan data [11]. Hasil dari penelitian ini berupa pengelompokan data sekolah serta universitas berdasarkan hasil perhitungan dengan menggunakan algoritma *K-Means*.

3. HASIL DAN PEMBAHASAN

Data yang diolah dalam penelitian ini merupakan sampel yang diambil dari mahasiswa Kabupaten Magetan yang sedang berkuliah di Universitas di Surabaya untuk Angkatan 2023 dan 2024.

3.1 Data Selection

Pada tahap ini, total sampel data memiliki jumlah sebanyak 250 data. Serta, akan dipilih hanya 3 atribut yang akan digunakan dalam penelitian yakni nama, asal sekolah, dan universitas.

Tabel 1. Data Mahasiswa yang Diperoleh

No .	Nama Lengkap	Asal Sekolah	Perguruan Tinggi
1.	Adista Larasati	Sman 3 Magetan	Upn Veteran Surabaya Jawa Timur
2.	Adista Larasati	Sman 3 Magetan	Upn Veteran Surabaya Jawa Timur

No .	Nama Lengkap	Asal Sekolah	Perguruan Tinggi
3.	Eri Eka Nugraheni	Sma Negeri 3 Magetan	Universitas Negeri Surabaya
...			
...			
...			
22 3.	Tya Ayu Devinta	Sma Negeri 1 Parang	Universitas Islam Negeri Sunan Ampel Surabaya
22 4.	Faiz Armathori Firmansyah	Sman 3 Magetan	Universitas Negeri Surabaya
22 5.	Faiz Armathori Firmansyah	Sman 3 Magetan	Universitas Negeri Surabaya

3.2 Pre-processing

Tahap *pre-processing* ini dilakukan dalam beberapa tahap diantaranya *data integration* dan *data cleaning*. Pada tahap *data integration*, kedua data yang sebelumnya terpisah disatukan menjadi satu dataset. Selain itu, pada tahap ini data disamakan formatnya, baik penulisan maupun aspek penulisan lainnya. Pada tahapan ini, diperoleh total data sebanyak 250 data. kemudian dilakukanlah *data cleaning* atau pembersihan data. Setiap ada kolom yang kosong pada suatu baris, ataupun data duplikasi, maka baris tersebut akan dihapus. Data setelah dibersihkan berubah menjadi 225 data total.

3.3 Transformation

Transformasi data dilakukan supaya dataset dapat diproses menggunakan metode *K-means Clustering*. Setiap data dalam ketiga atribut akan diubah dari semulanya kualitatif menjadi data numerik kuantitatif. Untuk atribut nama, setiap nama mahasiswa akan ditransformasi dalam bentuk format M1,M2,... dan seterusnya. Untuk asal sekolah, akan dibagi menjadi 3 kelompok yakni SMA ditransformasikan dengan nilai 1, SMK ditransformasikan dengan nilai 2, dan asal sekolah selain SMA dan SMK ditransformasikan dengan nilai 3. Sedangkan untuk universitas akan ditransformasi dalam 11 kelompok, diantaranya yakni UPN bernilai 1, UNESA bernilai 2, UNAIR bernilai 3, UINSA bernilai 4, PENS bernilai 5, ITS bernilai 6, PPNS bernilai 7, Poltekkes bernilai 8, Telkom bernilai 9, dan Muhammadiyah bernilai 10. Adapun hasil transformasi dapat dilihat melalui tabel berikut.

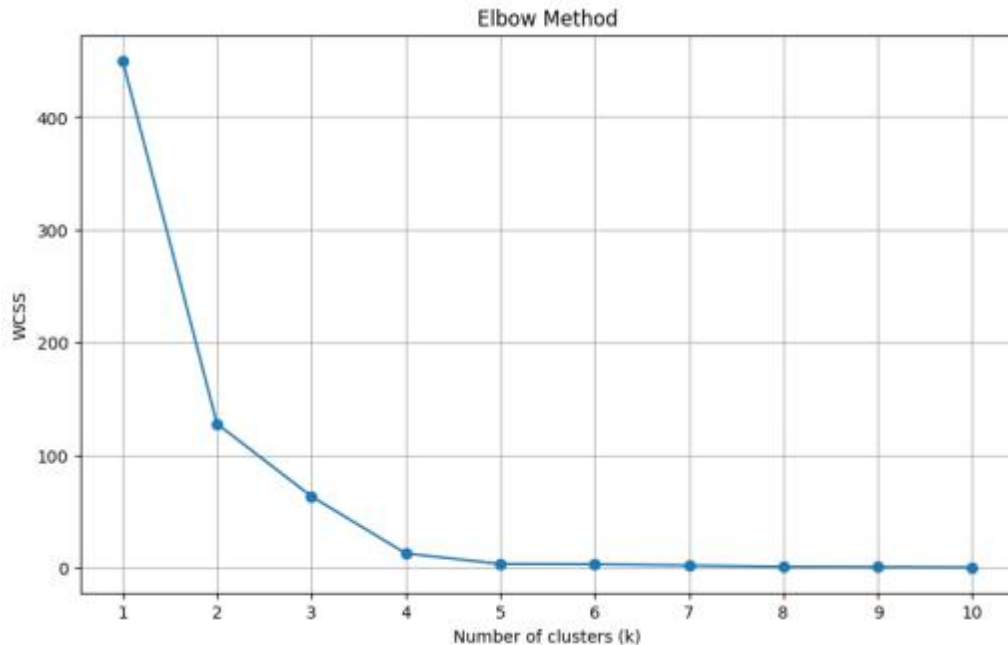
Tabel 2. Hasil Transformasi Data

Nama	Sekolah	Universitas
M1	1	1
M2	1	2
M3	1	3
....		
....		
....		
M223	3	4
M224	1	4
M225	1	2

3.4 Proses Training

Penelitian kali ini menggunakan *teknik clustering* dengan algoritma *K-Means*. Serta memilih jumlah cluster yang akan digunakan dengan membandingkan hasil metode *Elbow* dan *Silhouette Coefficient*. Berikut adalah langkah-langkah nya :

- 1) Menentukan K sebagai jumlah cluster dengan menggunakan Metode *Elbow*



Gambar 2. Hasil Pengujian Metode *Elbow*

Berdasarkan gambar 2 hasil pengujian metode elbow, maka jumlah *cluster* yang baik digunakan adalah 4 *cluster*, sehingga dalam penelitian ini menggunakan 4 *cluster* yaitu *cluster* 0, *cluster* 1, *cluster* 2, *cluster* 3.

- 2) Menentukan K sebagai jumlah *cluster* dengan Metode *Silhouette Coefficient*

```
Jumlah Cluster: 2 -> Silhouette Coefficient rata-rata: 0.62
Jumlah Cluster: 3 -> Silhouette Coefficient rata-rata: 0.58
Jumlah Cluster: 4 -> Silhouette Coefficient rata-rata: 0.56
Jumlah Cluster: 5 -> Silhouette Coefficient rata-rata: 0.55
Jumlah Cluster: 6 -> Silhouette Coefficient rata-rata: 0.54
Jumlah Cluster: 7 -> Silhouette Coefficient rata-rata: 0.53
Jumlah Cluster: 8 -> Silhouette Coefficient rata-rata: 0.53
Jumlah Cluster: 9 -> Silhouette Coefficient rata-rata: 0.52
```

Gambar 3. Hasil Pengujian Metode *Silhouette Coefficient*

Berdasarkan gambar 3 hasil pengujian metode silhouette coefficient, maka jumlah *cluster* yang baik digunakan adalah 2 *cluster*.

- 3) Berdasarkan analisis perbandingan antara metode *Elbow* yang menunjukkan titik optimal di $k=4$ dan *Silhouette Coefficient* yang menunjukkan nilai tertinggi 0.62 pada $k=2$, pemilihan 2 *cluster* menjadi lebih optimal karena memiliki kohesivitas data yang lebih baik dan model yang lebih sederhana.
- 4) Menghitung jarak setiap data antara *cluster*

	Nama	Sekolah	Universitas	Cluster_1_Distance	Cluster_2_Distance
0	1	3	1	2.837128	5.227400
1	2	3	1	2.834495	5.220866
2	3	3	1	2.831943	5.214369
3	4	3	1	2.829473	5.207910
4	5	3	1	2.827084	5.201488
...	...	...	...	...	...
220	221	2	10	3.733745	1.856169
221	222	2	10	3.745693	1.865850
222	223	2	10	3.757666	1.875606
223	224	2	10	3.769664	1.885438
224	225	2	10	3.781687	1.895344

Gambar 4. Menghitung Setiap Data Antara *Cluster*

- 5) Mengklasifikasikan setiap data berdasarkan kedekatannya dengan *centroid* (jarak terkecil).

	Nama	Sekolah	Universitas	Cluster
0	1	3	1	1
1	2	3	1	1
2	3	3	1	1
3	4	3	1	1
4	5	3	1	1
220	221	2	10	2
221	222	2	10	2
222	223	2	10	2
223	224	2	10	2
224	225	2	10	2

Gambar 5. Menentukan *Cluster* yang Paling Dekat

- 6) Melakukan perulangan dari langkah ke-2 hingga ke-4 sampai anggota tiap *cluster* tidak ada yang berubah.

Tabel 3. Hasil *Cluster* Baru Epoch Ke-1

Cluster	Nama	Sekolah	Universitas
1	33.0	1.646154	4.907692
2	145.5	1.131250	9.137500

- 7) Setelah melakukan perulangan sampai *epoch* ke-5 Hasil *cluster* setelah *epoch* ke-5 adalah sebagai berikut.

Tabel 4. Hasil Data dari *Epoch* Ke-5 dengan Dua *Cluster*

Cluster	Universitas	Sekolah
1	UPN : 10	Lainnya : 10
	UNESA : 4	Lainnya : 4
	UNAIR : 1	Lainnya : 1
	UINSA : 3	Lainnya : 3
	PENS : 2	Lainnya : 2
	ITS : 39	SMA : 38
	PPNS : 6	Lainnya : 1
2	Poltekkes : 1	SMA : 6
	Telkom : 136	SMA : 1
	Muhammadiyah : 23	SMA : 136
		SMA : 2
		SMK : 21

Selain itu percobaan dilakukan dengan menggunakan 3 *cluster* dan 4 *cluster*. Berikut merupakan hasil dari percobaan 3 *cluster* dan 4 *cluster* :

Tabel 5. Hasil Data dengan Tiga *Cluster*

Universitas	Sekolah
Hasil Cluster 1	
Telkom : 136	SMA : 136
Muhammadiyah : 23	SMA : 2
	SMK : 21
Hasil Cluster 2	
ITS : 28	SMA : 38
PPNS : 6	SMA : 6
Poltekkes : 1	SMA : 1
Hasil Cluster 3	
UPN : 10	Lainnya : 10
UNESA : 4	Lainnya : 4
UNAIR : 1	Lainnya : 1
UINSA : 3	Lainnya : 3
PENS : 2	Lainnya : 2
ITS : 1	Lainnya : 1

Tabel 6. Hasil Data dengan Empat *Cluster*

Universitas	Sekolah
Hasil Cluster 1	
Telkom : 136	SMA : 136
Muhammadiyah : 2	SMA : 2
Hasil Cluster 2	
ITS : 38	SMA : 38
PPNS : 6	SMA : 6
Poltekkes : 1	SMA : 1
Hasil Cluster 3	
UPN : 10	Lainnya : 10
UNESA : 4	Lainnya : 4
UNAIR : 1	Lainnya : 1
UINSA : 3	Lainnya : 3
PENS : 2	Lainnya : 2
ITS : 1	Lainnya : 1
Hasil Cluster 4	
Muhammadiyah : 21	SMK : 21

Berdasarkan hasil, percobaan yang sudah dilakukan dengan 2 *cluster* menunjukkan proses pengelompokan lebih mudah dilakukan, daripada dengan 3 atau 4 *cluster*. Selain itu 2 *cluster* memiliki nilai *silhouette* tertinggi (0.62) dari *cluster* lain. Meskipun 3 *cluster* dan 4 *cluster* memberikan lebih banyak *cluster*, hasil pengelompokannya justru kurang bermakna dan cenderung memisahkan data yang seharusnya bisa dikelompokkan bersama. 2 *cluster* memberikan hasil yang lebih masuk akal dan lebih mudah diinterpretasikan untuk analisis lebih lanjut.

4. KESIMPULAN DAN SARAN

4.1 Kesimpulan

Hasil pembahasan *clustering* data mahasiswa Kabupaten Magetan yang sedang berkuliah di Universitas di Surabaya menggunakan metode *K-Means* menunjukkan bahwa berdasarkan hasil cluster dataset dari mahasiswa angkatan 2023 dan 2024 memiliki perbedaan hasil ketika menggunakan metode *elbow* dengan metode *silhouette*. Berdasarkan perhitungan *elbow*, didapatkan jumlah *cluster* yang optimal yakni sebanyak 4 *cluster*. Sedangkan pada metode *silhouette* jumlah *cluster* yang baik digunakan ialah 2 *cluster* dengan nilai tertingginya 0,62. Dengan 2 *cluster* menjadikan perhitungan lebih optimal, karena kohesivitas yang lebih baik serta model yang lebih sederhana. Hasil akhir dari perhitungannya ialah pada *epoch* ke 5 dengan *cluster* 1 sebanyak 65 *items* dan *cluster* 2 dengan 160 *items*. Dengan adanya data ini nantinya dapat dimanfaatkan sebagai strategi promosi bagi calon mahasiswa yang ada di Kabupaten Magetan berdasarkan *cluster* yang telah terbentuk.

4.2 Saran

Adapun saran untuk penelitian berikutnya, diantaranya :

1. Pengelompokan data mahasiswa di Kabupaten Magetan yang ada di Surabaya ini sebaiknya dilakukan setiap tahunnya, supaya mendapatkan hasil yang lebih akurat.
2. Penelitian ini dapat dijadikan salah satu referensi untuk yang akan mendatang.
3. Diharapkan untuk menambahkan ataupun mengkombinasikan metode *clustering* yang lainnya pada penelitian berikutnya untuk menghasilkan penelitian yang lebih baik.

5. DAFTAR RUJUKAN

- [1] F. R. Hariri and R. A. Ramadhani, "Seminar Nasional Teknologi Informasi, Komunikasi dan Aplikasinya Volume 04, Tahun 2017".
- [2] T. Alfina, B. Santosa, and J. A. R. Hakim, "Analisa Perbandingan Metode Hierarchical Clustering, K-means dan Gabungan Keduanya dalam Cluster Data (Studi kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS)," vol. 1, 2012.
- [3] W. Lestari, "Clustering Data Mahasiswa Menggunakan Algoritma K-Means Untuk Menunjang Strategi Promosi (Studi Kasus : STMIK Bina Bangsa Kendari)," *SIMKOM*, vol. 4, no. 2, pp. 35–48, Jul. 2019, doi: 10.51717/simkom.v4i2.37.
- [4] U. W. Latifah, S. Bahri, and M. Satriandhini, "Implementasi Algoritma K-Means Clustering untuk Strategi Promosi Kampus IBISA," *JIKO J. Inform. Dan Komput.*, vol. 8, no. 2, p. 292, Sep. 2024, doi: 10.26798/jiko.v8i2.1307.
- [5] R. Julianto Dwi Putra, F. Tri Anggraeny, and Sugiarto, "Analisis Kualitas Website Fakultas Hukum Universitas Hang Tuah Dengan WebQual 4.0 Dan IPA," *J. Inform. Dan Sist. Inf.*, vol. 2, no. 2, pp. 322–333, Jul. 2021, doi: 10.33005/jifosi.v2i2.300.
- [6] M. H. Adiya and Y. Desnelita, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru," vol. 05, no. 01, 2019.
- [7] Ramadhana, Islamiyah, and A. P. A. Masa, "Penerapan Data Mining Menggunakan Metode K-Means Clustering Pada Data Ekspor Batubara," *Adopsi Teknol. Dan Sist. Inf. ATASI*, vol. 2, no. 1, pp. 35–42, Jun. 2023, doi: 10.30872/atasi.v2i1.595.
- [8] J. Faran and R. T. Aldisa, "Perbandingan Algoritma K-Means dan K-Medoids Dalam Pengelompokan Kelas Untuk Mahasiswa Baru Program Magister," vol. 5, no. 2, 2024.
- [9] A. Nursikuwagus, Suherman, and I. Alamsyah, "KLUSTERISASI PENYEBAB KEMATIAN DI INDONESIA DENGAN PENERAPAN ALGORITMA K-MEANS," *JSiIJ. Sist. Inf.*, vol. 11, no. 2, pp. 56–63, Sep. 2024, doi: 10.30656/jsii.v11i2.9276.
- [10] H. S. Pakpahan, J. A. Widians, and H. D. A. Firmanda, "Implementasi Metode K-Means Untuk Pengelompokan Potensi Produksi Komoditas Perkebunan," *Adopsi*

Teknol. Dan Sist. Inf. ATASI, vol. 1, no. 1, pp. 52–60, Jun. 2022, doi: 10.30872/atasi.v1i1.49.

- [11] A. Yudhistira and R. Andika, “Pengelompokan Data Nilai Siswa Menggunakan Metode K-Means Clustering,” *J. Artif. Intell. Technol. Inf. JAITI*, vol. 1, no. 1, pp. 20–28, Feb. 2023, doi: 10.58602/jaiti.v1i1.22.