

Klasterisasi Data Travel Umroh di Marketplace Umroh.com Menggunakan Metode K-Means

Mas'ud Hermansyah¹, Rifky Aditia Hamdan², Fazar Sidik³, Arief Wibowo⁴

Program Studi Magister Ilmu Komputer, Universitas Budi Luhur
Jl.Raya Ciledug, Petukangan Utara, Jakarta Selatan 12260 Indonesia

¹masudhermansyah@gmail.com

²rifkyaditia7@gmail.com

³fajarsidik1430@gmail.com

⁴arief.wibowo@budiluhur.ac.id

Abstrak

Perjalanan wisata religi khususnya untuk ibadah umroh semakin diminati masyarakat. Umroh.com adalah jembatan antara calon jamaah umroh dengan travel umroh untuk bertransaksi didalam persiapan perjalanan ibadah umroh. Umroh.com sebuah perusahaan berbentuk *marketplace* dengan menawarkan produk paket umroh dan menyediakan travel umroh lebih dari 60 travel umroh terpercaya di Indonesia dengan izin resmi Kementerian Agama. Untuk dapat melakukan persaingan antar travel umroh, masing-masing travel membutuhkan strategi-strategi penjualan untuk dapat menarik minat calon jamaah umroh dan meningkatkan laba atau pendapatan perusahaan travel. Pada penelitian ini analisis *data mining* dilakukan dengan Teknik *Clustering* menggunakan metode *K-Means*. Hasil akhir dari penelitian ini adalah mengelompokkan data travel yang terdaftar di Umroh.com untuk mengetahui data travel yang memiliki potensi atau kecenderungan calon jamaah umroh dalam memilih travel umroh tersebut. Hasil penelitian menunjukkan bahwa pemodelan *K-Means* dengan jumlah klaster sebanyak 5 klaster memiliki nilai DBI terkecil yaitu sebesar 0,134.

Kata Kunci: *Klasterisasi, Travel Umroh, Marketplace, K-Means, Davies Bouldin Index*

Abstract

Religious tourism, especially for Umrah worship, is increasingly in demand by the public. Umroh.com is a bridge between prospective Umrah pilgrims and Umrah travel to make transactions in preparation for the Umrah pilgrimage trip. Umroh.com is a company in the form of a marketplace by offering Umrah package products and providing Umrah travel for more than 60 trusted Umrah travel agents in Indonesia with official permission from the Ministry of Religion. To be able to compete between Umrah travels, each travel needs sales strategies to attract prospective Umrah pilgrims and increase the profit or income of the travel company. In this study, data mining analysis was carried out by using the Clustering Technique using the K-Means method. The final result of this research is to classify the travel data listed on Umroh.com to find out which travel data has the potential or tendency of prospective Umrah pilgrims to choose the Umrah travel. The results showed that the K-Means modeling with the number of clusters of 5 clusters had the smallest DBI value of 0.134.

Keywords: Clusterization, Umrah Travel, Marketplace, K-Means, Davies Bouldin Index

1. Pendahuluan

Teknologi *data mining* di dalam sebuah perusahaan digunakan untuk mempelajari bagaimana memanfaatkan data-data historis dengan menemukan suatu pola-pola yang pada dasarnya digunakan agar bisa membantu mempercepat dalam proses pengambilan keputusan secara tepat dan memungkinkan perusahaan untuk mengelola informasi menjadi sebuah pengetahuan yang baru. Melalui pengetahuan baru yang didapat, perusahaan diharapkan bisa meningkatkan pendapatannya, dan pada akhirnya dimasa yang akan datang perusahaan dapat lebih kompetitif [1].

Umroh.com adalah jembatan antara calon jamaah umroh dengan travel umroh untuk bertransaksi didalam perjalanan ibadah umroh. Umroh.com merupakan perusahaan swasta berbentuk *marketplace* dengan melayani penjualan produk paket umroh dan menyediakan paket umroh lebih dari 60 travel umroh terpercaya di Indonesia dengan izin resmi Kementerian Agama (Kemenag). Persiapan ibadah umroh akan terasa mudah karena calon jamaah umroh tidak lagi memusingkan untuk mencari travel umroh dan menemukan paket umroh yang diinginkan.

Meningkatnya jumlah data jamaah umroh pada penyedia layanan travel umroh tentunya memiliki variasi strategi yang berbeda-beda. Pengelompokan data travel umroh didalam *marketplace* Umroh.com merupakan salah satu tujuan untuk meningkatkan kualitas pelayanan dari masing-masing penyedia layanan travel umroh. Pengelompokan ini bertujuan untuk mengetahui minat calon jamaah umroh pada Umroh.com berdasarkan: jumlah jamaah yang sudah diberangkatkan pada masing-masing travel pada tahun 2019, rata-rata harga paket pada masing-masing travel, dan banyaknya paket umroh yang tersedia pada masing-masing travel umroh.

Pengelompokan tersebut dapat menggunakan metode pengelompokan algoritma *K-Means Clustering*. *Clustering* merupakan metode dalam *data mining* dan algoritma *K-Means* dipilih karena kesederhanaan algoritma dan efisiensinya. Hasil dari mengelompokkan data travel yang dipilih oleh jamaah umroh adalah untuk mengetahui data yang memiliki potensi atau kecenderungan pelanggan dalam memilih travel umroh tersebut. Hasil ini dapat digunakan untuk memberi saran pertimbangan dalam menentukan strategi pemasaran kepada masing-masing travel umroh.

2. Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah metode *data mining clustering* dengan menggunakan algoritma *K-Means*.

2.1. Data Mining

Data mining merupakan salah satu bidang paling penting dalam penelitian yang bertujuan untuk memperoleh informasi dari data set. *Data mining* mulai ada sejak 1990-an sebagai cara yang efektif untuk mengambil pola dan informasi yang sebelumnya tidak diketahui dari suatu data set. Teknik *data mining* digunakan untuk menemukan hubungan antara data untuk melakukan pengklasifikasian yang digunakan dalam klasifikasi nilai-nilai dari beberapa variabel, membagi data yang diketahui menjadi kelompok-kelompok yang mempunyai kesamaan karakteristik (*clustering*). *Data mining* merupakan bagian dari proses penemuan pengetahuan dari basis data *Knowledge Discovery in Databases* (KDD), yang mana merupakan tahapan dari KDD [2].

2.2. Klasterisasi

Clustering merupakan sebuah teknik pemrosesan data yang digunakan untuk menemukan pola-pola tersembunyi pada kumpulan data [3]. *Clustering* atau klasterisasi adalah suatu alat bantu pada *data mining* yang bertujuan untuk mengelompokkan objek-objek ke dalam beberapa klaster. Klaster adalah sekumpulan objek-objek data yang memiliki kemiripan karakteristik satu sama lain dalam klaster yang sama dan berbeda karakteristik terhadap objek-objek yang berbeda klaster [4].

Cara kerja teknik ini adalah dengan mengelompokkan sekumpulan data ke dalam kelas-kelas atau kluster-kluster, yang mana objek-objek yang ada pada kelas tersebut memiliki similaritas yang tinggi jika dibandingkan dengan objek lain yang ada dalam kelas tersebut, namun memiliki similaritas yang rendah jika dibandingkan dengan objek yang ada di kelas atau kluster lain [5].

2.3. Algoritma *K-Means*

Algoritma *K-Means* adalah metode yang termasuk dalam algoritma *clustering* berbasis jarak yang membagi data ke dalam sejumlah *cluster* dan algoritma ini hanya bekerja pada atribut numerik [6].

Metode dasar analisis algoritme *K-Means Clustering* adalah sebagai berikut:

- a. Tentukan jumlah klaster (*k*), tetapkan pusat *cluster* secara acak.
- b. Hitung jarak setiap data ke pusat *cluster*.
- c. Kelompokan data ke dalam *cluster* dengan jarak yang paling pendek.

- d. Hitung pusat *cluster* baru.
- e. Ulangi langkah 2 (dua) sampai 4 (empat) apabila masih ada data yang berpindah kelompok atau apabila ada perubahan nilai *centroid* di atas nilai ambang yang ditentukan.

Proses *clustering* dimulai dengan mengidentifikasi data yang diklasterisasi, dapat digunakan rumus formula *Euclidean Distance* seperti yang terlihat pada rumus persamaan (1) [5]. Untuk menghitung jarak semua data ke setiap titik pusat *cluster* dapat menggunakan teori jarak *Euclidean* yang dirumuskan sebagai berikut [7].

$$d_{ij} = \sqrt{\sum_{k=1}^m X_{ij} - C_{jk}^2} \tag{1}$$

Dimana d_{ij} merupakan jarak objek antar nilai data dan nilai pusat *cluster*, m adalah jumlah dimensi data, X_{ij} merupakan nilai data dari dimensi ke- k dan C_{jk} adalah nilai pusat *cluster* dari dimensi ke- k [5].

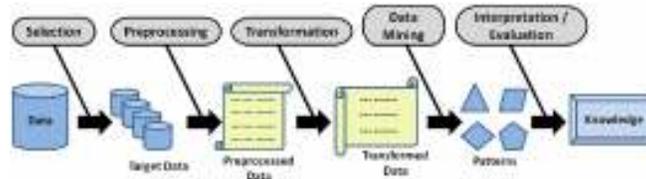
Untuk menghitung *centroid* baru, dapat menggunakan rumus persamaan (2), sebagai berikut:

$$C = \frac{\sum m}{n} \tag{2}$$

Dimana C merupakan *centroid* data, m adalah anggota data yang termasuk ke dalam *centroid* tertentu dan n adalah jumlah data yang menjadi anggota *centroid* tertentu [5].

2.4. Knowledge Discovery in Database (KDD)

KDD adalah metode yang digunakan untuk dapat memperoleh pengetahuan yang berasal dari *database* yang ada. Hasil pengetahuan yang diperoleh dapat dimanfaatkan untuk basis pengetahuan (*knowledge base*) yang digunakan dalam keperluan mengambil keputusan. Secara lebih detail, proses KDD seperti pada gambar berikut ini yang diadopsi dari [8]:



Gambar 1. Proses Knowledge Discovery in Database (KDD)

- a. *Data Selection*
Bertujuan mentransformasikan data mentah ke dalam format yang sesuai untuk bisa dilakukan analisa. Terdiri atas proses seleksi fitur, normalisasi dan *subsetting* data [5].
- b. *Data Preprocessing*
Pada *preprocessing* terdapat dua tahap, yaitu sebagai berikut:
 - 1. *Data Cleaning*
Menghilangkan data yang tidak diperlukan seperti menangani *missing value*, *noise* data serta menangani data-data yang tidak konsisten dan relevan.
 - 2. *Data Integration*
Dilakukan terhadap atribut yang mengidentifikasi entitas yang unik [8].
- c. *Data Transformation*
Transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses transformasi dalam KDD merupakan proses yang sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data [9].
- d. Proses *Mining*
Proses *mining* bertujuan untuk memberikan solusi kepada pengambil keputusan dalam bisnis guna meningkatkan kualitas bisnis perusahaan [10].
- e. *Interpretation / Evaluation*

Model informasi yang nantinya dihasilkan dari proses *mining* harus ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang bersangkutan. Pada tahap ini mencakup pemeriksaan apakah model atau informasi yang ditemukan dari proses data mining bertentangan dengan fakta atau hipotesis yang ada sebelumnya [11].

f. *Knowledge*

Pola-pola yang dihasilkan akan dipresentasikan kepada pengguna. Pada tahapan ini pengetahuan baru yang dihasilkan bisa dipahami semua orang yang akan dijadikan acuan pengambilan keputusan [8].

2.5. Davies Bouldin Index

Davies Bouldin Index (DBI) adalah menghitung rata-rata nilai setiap titik pada himpunan data. Perhitungan nilai setiap titik merupakan jumlah nilai *compactness* yang dibagi dengan jarak antara kedua titik pusat *cluster* sebagai *separation* [12].

Dalam proses penilaian dari model yang dihasilkan, digunakan *Davies Bouldin Index*. DBI digunakan untuk mengoptimalkan jarak diluar *cluster* dan meminimalkan jarak didalam *cluster* yang bisa dihitung dengan persamaan 3 berikut [5]:

$$S_i = \frac{1}{|c_i|} \sum_{x \in c_i} \{ |x - z_i| \} \quad (3)$$

Dimana c_i adalah banyaknya titik yang terdapat di dalam cluster i , x adalah data, dan z_i adalah *centroid* cluster i . Sedangkan jarak antar *cluster* didefinisikan pada Persamaan 4 berikut:

$$d_{ij} = |Z_i - Z_j| \quad (4)$$

Dimana z_i adalah *centroid* dari *cluster* i dan z_j adalah *centroid* dari *cluster* j . Perhitungan jarak d_{ij} dapat menggunakan *euclidean distance*. Selanjutnya mendefinisikan R_i , untuk *cluster* c_i pada persamaan 5 berikut:

$$R_{i,qt} = \max_{j, j \neq i} \left\{ \frac{S_{i,t} + S_{j,q}}{d_{ij,t}} \right\} \quad (5)$$

Selanjutnya *Davies Bouldin Index* didefinisikan pada persamaan 6 berikut:

$$DBI = \frac{1}{k} \sum_{i=1}^k R_{i,qt} \quad (6)$$

Dari persamaan 6, k merupakan jumlah *cluster* yang digunakan. Semakin kecil nilai yang diperoleh dari perhitungan DBI (non-negatif ≥ 0), maka memberikan hasil semakin baik pula *cluster* yang diperoleh dari pengelompokan *cluster* dengan metode *K-Means* yang digunakan [13].

3. Hasil dan Pembahasan

Langkah-langkah untuk melakukan *data mining* mengikuti aturan KDD adalah sebagai berikut:

a. *Data Selection*

Data travel yang diolah merupakan data tahun 2019 dan memiliki atribut-atribut sebagai berikut:

1. Nama Travel
2. Alamat
3. Jamaah Keberangkatan
4. Harga Paket Umroh
5. Jumlah Paket Umroh
6. Tahun Izin Lisensi Umroh
7. Tahun Izin Lisensi Haji
8. Nomor Telepon

Dari atribut yang telah disebutkan sebelumnya, dipilih hanya empat atribut. yaitu:

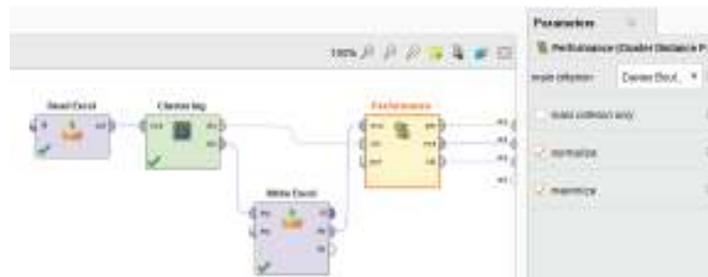
1. Nama Travel
2. Jumlah Keberangkatan

3. Harga Paket Umroh
 4. Jumlah Paket Umroh
- b. *Data Preprocessing / Data Cleaning*
Langkah ini dilakukan untuk membersihkan data duplikasi dan data yang tidak konsisten. Dalam penelitian ini proses *cleaning* adalah membersihkan nama-nama travel yang sama atau nama travel yang sudah tidak aktif.
- c. *Data Transformation*
Transformasi dilakukan untuk menyesuaikan data agar dapat diproses. Transformasi yang dilakukan adalah atribut Harga Paket Umroh dan kemudian ditransformasi menjadi rata-rata total Harga Paket Umroh.

Tabel 1. Data Travel Umroh

No	Nama Atribut	Tipe Data	Nilai
1	Nama Tarvel	Text	A – Z
2	Jumlah Keberangkatan	Integer	0 – 9
3	Rata-rata Harga Paket	Integer	0 – 9
4	Jumlah Paket	Integer	0 – 9

- d. *Proses Mining*
Proses *mining* dilakukan dengan tujuan untuk mencari informasi atau pola untuk *clustering* dengan menggunakan algoritma *K-Means*. Implementasi algoritma *K-Means* dalam penelitian ini menggunakan software *Rapidminer*. Pemodelan *K-Means Clustering* dengan *Rapidminer* dapat dilihat pada Gambar 2.



Gambar 2. Pemodelan *K-Means Clustering* dengan *Rapidminer*

- *Read Excel* adalah operator untuk membaca *ExampleSet* dari file Excel yang ditentukan.
- *Clustering* adalah operator yang melakukan proses pengelompokan menggunakan algoritma *K-Means*.
- *Write Excel* adalah operator untuk membuat laporan hasil *clustering* dengan tipe file .xlsx.
- *Performance* adalah operator yang digunakan untuk evaluasi kinerja metode *clustering* berdasarkan *centroid*.

Pada langkah ini dilakukan 5 kali percobaan proses pengklusteran dengan jumlah *cluster* mulai 2, 3, 4, 5, dan 6 *cluster*.

Tabel 2. Hasil Pemodelan *K-Means Clustering* dengan *Rapidminer*

Jumlah Cluster	Nama Cluster	Jumlah Anggota
2 cluster	Cluster 1	61
	Cluster 2	5

3 cluster	Cluter 1	56
	Cluster 2	1
	Cluster 3	9
4 cluster	Cluster 1	26
	Cluster 2	1
	Cluter 3	6
	Cluster 4	33
5 cluster	Cluster 1	29
	Cluter 2	4
	Cluster 3	11
	Cluter 4	1
	Cluster 5	21
6 cluster	Cluster 1	21
	Cluster 2	1
	Cluster 3	4
	Cluster 4	5
	Cluster 5	18
	Cluster 6	17

e. *Interpretation / Evaluation*

Langkah terakhir dari KDD ini dilakukan untuk menjelaskan makna dari setiap *cluster* yang terbentuk. Langkah ini dilakukan dengan cara membagi *centroid* dari iterasi terakhir dengan jumlah anggota setiap *cluster*-nya. Sehingga dapat disimpulkan deskripsi dari setiap *cluster*. Selain itu pada langkah ini juga melakukan evaluasi *cluster* dengan metode *Davies Bouldin Index* untuk menghitung rata-rata nilai setiap titik pada himpunan data [5].

Dilakukan beberapa kali percobaan terhadap model yang dibangun. Model pertama adalah melakukan *clustering* terhadap *dataset* tanpa melakukan *normalize*. Jumlah *cluster* ditentukan mulai dari 2, 3, 4, 5 dan 6. Hasil segmentasi yang terbentuk akan dievaluasi menggunakan *Davies Bouldin Index*. *Davies Bouldin Index* merupakan metode validasi *cluster* dari hasil *clustering*.

Tabel 3. Nilai *Davies Bouldin Index* Hasil Percobaan

Jumlah Cluster	Nilai DBI
2 cluster	0,146
3 cluster	0,147
4 cluster	0,136
5 cluster	0,134
6 cluster	0,135

Dari hasil percobaan yang dilakukan, algoritma *K-Means* dengan *Rapidminer*, jumlah 5 *cluster* menghasilkan kualitas *cluster* yang lebih baik dibandingkan dengan jumlah *cluster* 2, 3, 4, dan 6. Hasil evaluasi *cluster* menunjukkan bahwa algoritma *K-Means* dengan jumlah *cluster* 5 lebih optimal dengan nilai DBI paling kecil, sebesar 0.134. Jumlah 5 *cluster* yang terbentuk dengan segmentasi sebagai berikut:

Tabel 4. Hasil *Cluster*

Nama Cluster	Jumlah Anggota	Inisial Nama Travel
Cluster 1	29	AATOWA, AATR, ABDAAN, ADDBHUA, ALITVEL, ABAZHA, BMI, BTVL, BMW, DDTL, DNSTL, GABI, GMAWA, HASWD, KAJ, KIT, KOAPRNG, MDITRVL, MWSTA, MNRWSTA, MFRTUR,

		NAN, PTITHU, QBTWST, RHMTUR, SUMRAH, TIBITUR, TRZTRVEL, WINTUR
Cluster 2	4	ANSATRVL, CTL, PATRVL, STUR
Cluster 3	11	AHZ, AG, ADSTUR, AGGTP, FHY, GLSA, KZZHTUR, MTMUTM, PAZZWA, TNMTHH, TDTIO
Cluster 4	1	AANUL
Cluster 5	21	AHZ, AFATRS, AFHMAISI, AMRA, BCMI, BTITUR, CAW, FTHIH, GTTS, JJKIMI, JDHTR, MTTVEL, MBATUR, PNATUR, RBITUR, RSITUR, RTTRVL, STUR, SOLUTUR, TCUTMA, ZITTUR

Dari data tabel di atas yang terlihat bahwa *cluster 1* merupakan kelompok travel umroh dengan jumlah keberangkatan besar dengan nilai keberangkatan mencapai 8.000 jamaah, rata-rata harga paket tergolong murah, antara Rp. 17.000.000,00 – Rp. 25.000.000,00, rata-rata pilihan paket umroh cukup banyak, yaitu mencapai 30 paket umroh yang tersedia, dan tahun pendirian travel antara tahun 1988 sampai tahun 2017.

Cluster 2 merupakan kelompok travel umroh dengan jumlah keberangkatan rendah, yaitu dibawah 3.200 jamaah, dengan harga paket tergolong mahal yaitu dengan biaya sekitar Rp. 40.000.000,00, rata-rata pilihan paket umroh relatif banyak, yaitu 70 paket umroh yang tersedia, dan tahun pendirian travel antara tahun 1972 sampai tahun 2009.

Cluster 3 merupakan kelompok travel umroh dengan jumlah keberangkatan cukup rendah yaitu di bawah 2.500 jamaah, rata-rata harga paket tergolong mahal yaitu rata-rata sebesar Rp. 30.000.000,00 rata-rata pilihan paket umroh relatif banyak, yaitu 70 paket umroh yang tersedia, dan tahun pendirian travel antara tahun 1989 sampai tahun 2017.

Cluster 4 merupakan kelompok travel umroh dengan jumlah keberangkatan sangat rendah dengan jumlah keberangkatan dibawah 280 jamaah, rata-rata harga paket tergolong sangat mahal, yaitu mencapai Rp. 60.000.000,00, rata-rata pilihan paket umroh sangat sedikit, yaitu hanya 4 paket umroh yang tersedia, dan tahun pendirian travel tahun 1999.

Cluster 5 merupakan kelompok travel umroh dengan jumlah keberangkatan relatif tinggi dengan jumlah keberangkatan mencapai 6.000 jamaah, rata-rata harga paket tergolong relatif murah, yaitu antara Rp. 25.000.000,00, sampai Rp. 30.000.000,00, rata-rata pilihan paket umroh sanga banyak, yaitu 120 paket umroh yang tersedia, dan tahun pendirian travel antara tahun 1971 sampai tahun 2018.

4. Kesimpulan

Penerapan metode *K-Means* dalam pengelompokan data travel umroh pada *marketplace* Umroh.com dapat menghasilkan kelompok daftar travel yang paling diminati, cukup diminati dan kurang diminati. Hasil klasterisasi dapat menjadi rujukan bagi manajemen untuk mengatur strategi pemasaran kepada masing-masing travel umroh. Berdasarkan hasil pengelompokan diperoleh pengelompokan dengan 5 *cluster* merupakan hasil pengelompokan yang paling optimal dengan nilai DBI terkecil, yaitu 0,134. Kondisi ini merupakan representasi dari hasil pengolahan data pelanggan atau jamaah yang mendaftar pada web <http://www.umroh.com>.

Referensi

- [1] A. Sani, "Penerapan Metode K-Means Clustering Pada Perusahaan Penerapan Metode K- Means Clustering Pada Perusahaan," *ResearchGate*, no. August, p. 7, 2018.
- [2] Z. Mustofa and I. S. Suasana, "Algoritma Clustering K-Medoids Pada E-Government Bidang Information And Communication," *J. Teknol. dan Komun.*, vol. 9, pp. 1–10, 2018.
- [3] I. A. Musdar and A. SN, "Metode RCE-Kmeans untuk Clustering Data," *IJCCS*, vol. 9, no. 2, pp. 157–166, 2015.
- [4] N. R. W, S. Defiyanti, and M. Jajuli, "Implementasi Algoritma K-Means dalam Pengklasifikasian Mahasiswa Pelamar Beasiswa," *J. Ilm. Teknol. Inf. Terap.*, vol. 1, no. 2, pp. 62–68, 2015.
- [5] I. R. Mahartika and A. Wibowo, "Data Mining Klasterisasi dengan Algoritme K-Means untuk Pengelompokan Provinsi Berdasarkan Konsumsi Bahan Bakar Minyak Nasional," in *Prosiding Seminar Nasional Sisfotek*, 2019.
- [6] W. M. P. Dhuhita, "Clustering Menggunakan Metode K-Means Untuk Menentukan Status Gizi Balita," *J. Inform.*, vol. 15, no. 2, pp. 160–174, 2015.
- [7] R. Sibarani and Chafid, "Algorithma K-Means Clustering Strategi Pemasaran Penerimaan Mahasiswa Baru Universitas Satya Negara Indonesia [Algorithma K-Means Clustering Strategy Marketing Admission Universitas Satya Negara Indonesia]," *Semin. Nas. Cendekiawan*, vol. 4, pp. 685–690, 2018.
- [8] M. H. Adiya and Y. Desnelita, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru," *J. Nas. Teknol. dan Sist. Inf.*, vol. 01, pp. 17–24, 2019.
- [9] Y. Asriningtias and R. Mardhiyah, "Aplikasi Data Mining untuk Menampilkan Informasi Tingkat Kelulusan Mahasiswa," *J. Inform.*, vol. 8, no. 1, pp. 837–848, 2014.
- [10] M. H. Siregar, "Klasterisasi Penjualan Alat-Alat Bangunan Menggunakanmetode K-Means (Studi Kasus Di Toko Adi Bangunan)," *J. Teknol. dan Open Source*, vol. 1, no. 2, pp. 83–91, 2018.
- [11] D. Mukhibah and A. Kurniawati, "Implementasi Data Mining dalam Prediksi Performance Software Engineeri PT. Emerio Menggunakan Decision Tree," *J. Inform. dan Komput.*, vol. 22, no. 1, pp. 31–43, 2017.
- [12] A. F. Khairati, A. A. Adlina, G. F. Hertono, and B. D. Handari, "Kajian Indeks Validitas pada Algoritma K-Means Enhanced dan K-Means MMCA," in *Prosiding Seminar Nasional Matematika*, 2019, vol. 2, pp. 161–170.
- [13] F. Ramdhan, A. Hoyyi, and Moch Abdul Mukid, "Pengelompokan Provinsi di Indonesia Berdasarkan Karakteristik Kesejahteraan Rakyat Menggunakan Metode K-Means Cluster," *J. Gaussian*, vol. 4, no. 2006, pp. 875–884, 2015.