

ALGORITMA K-MEANS UNTUK KLASTERISASI JUMLAH JEMAAH HAJI YANG DIBERANGKATKAN KE TANAH SUCI MEKAH MENURUT KABUPATEN/KOTA DI PROVINSI SUMATERA UTARA

M. Fachrurrozi Nasution¹⁾, Labuan Nababan²⁾, Ernida Marbun³⁾, Lamtiur Sinambela⁴⁾, Dedek Indra Gunawan Hts⁵⁾

^{1,2,3,5)} Universitas Satya Terra Bhinneka

⁴⁾ Politeknik Negeri Medan

Email : fachrurrozi@satyaterabhinneka.ac.id¹⁾, labuannababan@satyaterabhinneka.ac.id²⁾, ernidamarbun@satyaterabhinneka.ac.id³⁾, lamtiursinambela@polmed.ac.id⁴⁾, dedekindra@satyaterabhinneka.ac.id⁵⁾

ABSTRAK

Setiap tahun, Provinsi Sumatera Utara memberangkatkan ribuan jemaah haji ke Tanah Suci Mekah. Distribusi jumlah jemaah yang berasal dari berbagai kabupaten/kota menunjukkan pola yang beragam dan perlu dianalisis lebih lanjut guna mendukung perencanaan dan pengambilan keputusan oleh instansi terkait. Penelitian ini bertujuan untuk melakukan klasterisasi terhadap jumlah jemaah haji berdasarkan wilayah kabupaten/kota di Provinsi Sumatera Utara menggunakan algoritma K-Means, salah satu metode unsupervised learning yang efektif untuk pengelompokan data. Data yang digunakan dalam penelitian ini merupakan data sekunder jumlah jemaah haji dari setiap kabupaten/kota yang diperoleh dari sumber resmi. Proses klasterisasi dilakukan melalui beberapa tahapan, yaitu inisialisasi centroid awal, penghitungan jarak Euclidean, pengelompokan data berdasarkan kedekatan terhadap centroid, dan pembaruan posisi centroid hingga proses konvergen. Untuk menentukan jumlah klaster yang optimal, digunakan metode Elbow. Hasil analisis menunjukkan bahwa data dapat dibagi ke dalam tiga klaster utama, yaitu klaster dengan jumlah jemaah tinggi, sedang, dan rendah. Visualisasi klaster membantu memperlihatkan distribusi spasial yang lebih jelas dari masing-masing kelompok wilayah. Implementasi algoritma K-Means dalam penelitian ini membuktikan bahwa metode klasterisasi dapat digunakan secara efektif dalam menganalisis data demografis, khususnya dalam konteks pelayanan ibadah haji. Seluruh proses analisis dilakukan menggunakan Google Colab, yang memfasilitasi komputasi berbasis cloud dan mendukung visualisasi data secara interaktif.

Kata kunci: Klasterisasi, K-Means, Jemaah Haji, Sumatera Utara, Data Mining

ABSTRACT

Every year, North Sumatra Province sends thousands of pilgrims to the Holy Land of Mecca. The distribution of pilgrims from various districts and cities shows diverse patterns that need to be analyzed to support planning and decision-making by relevant authorities. This study aims to perform clustering on the number of Hajj pilgrims based on district/city-level data in North Sumatra Province using the K-Means algorithm, one of the effective unsupervised learning methods for data grouping. The data used in this study are secondary data on the number of Hajj pilgrims from each district/city, obtained from official sources. The clustering process involves several stages, including initial centroid selection, Euclidean distance calculation, data grouping based on proximity to centroids, and centroid updates until convergence is reached. The Elbow Method was applied to determine the optimal number of clusters. The analysis results show that the data can be grouped into three main clusters, namely regions with high, medium, and low numbers of pilgrims. Cluster visualization further illustrates the spatial distribution of each regional group more clearly. The implementation of the K-Means algorithm in this study demonstrates that clustering methods can be effectively used to analyze demographic data, particularly in the context of Hajj service management. All data processing

and analysis were conducted using Google Colab, which facilitates cloud-based computation and supports interactive data visualization.

Keywords: Clustering, K-Means, Hajj Pilgrims, North Sumatra, Data Mining

1. PENDAHULUAN

Ibadah haji merupakan salah satu ibadah keagamaan terbesar di dunia yang melibatkan jutaan jamaah dari berbagai negara setiap tahunnya (Nusa et al., n.d.). Melaksanakan Ibadah haji bagi yang mampu merupakan rukun islam yang kelima bagi umat islam. Setiap tahun, Provinsi Sumatera Utara memberangkatkan ribuan jemaah haji ke Tanah Suci Mekah. Jumlah jemaah ini bervariasi antar kabupaten/kota, bergantung pada faktor-faktor seperti jumlah penduduk, tingkat ekonomi, dan kesadaran religius masyarakat (Hidayat, 2022). Oleh karena itu, diperlukan analisis yang dapat mengelompokkan kabupaten/kota berdasarkan jumlah jemaah haji yang diberangkatkan untuk mendukung perencanaan dan pengelolaan pemberangkatan haji yang lebih efektif.

Machine learning merupakan cabang kecerdasan buatan yang berfokus pada pembuatan sistem yang mampu belajar dan membuat keputusan berdasarkan data tanpa diprogram secara eksplisit. Dalam machine learning, terdapat dua jenis pembelajaran utama, yaitu supervised learning dan unsupervised learning (Nasution & Roesnelly, 2024). Pengelompokan wilayah berdasarkan jumlah jemaah haji dapat dilakukan menggunakan pendekatan analisis data, khususnya metode *clustering* (Wilayah et al., n.d.). Salah satu algoritma yang banyak digunakan dalam *unsupervised learning* adalah *K-Means Clustering*. Algoritma ini mampu mengelompokkan data ke dalam beberapa kluster berdasarkan kemiripan karakteristik, dalam hal ini adalah jumlah jemaah haji dari tiap kabupaten/kota (Humaira Mufidah et al., 2023). Proses klusterisasi ini dapat menghasilkan pola distribusi yang lebih jelas dan dapat menjadi landasan dalam perumusan kebijakan, perencanaan kuota, serta peningkatan pelayanan haji secara lebih merata dan efisien.

Algoritma K-Means adalah salah satu teknik klusterisasi paling populer yang digunakan untuk membagi data ke dalam kelompok-kelompok berdasarkan kemiripan karakteristik. Proses K-Means dimulai dengan pemilihan sejumlah pusat kluster (centroid) secara acak, dilanjutkan dengan pengelompokan data berdasarkan jarak terdekat ke centroid, dan pembaruan posisi centroid secara iteratif hingga mencapai konvergensi (Zhao & Zhou, 2021).

Dalam Analisis kluster teknik yang digunakan untuk mengelompokkan suatu objek pengamatan sehingga setiap kluster menunjukkan kesamaan dengan subjek yang diteliti. Beberapa metode dapat digunakan untuk melakukan clustering, diantaranya adalah metode k-means clustering. Dengan metode ini, sejumlah besar data dapat dikelompokkan dengan cepat dan efisien (Dakhi & Ningsi, 2024).

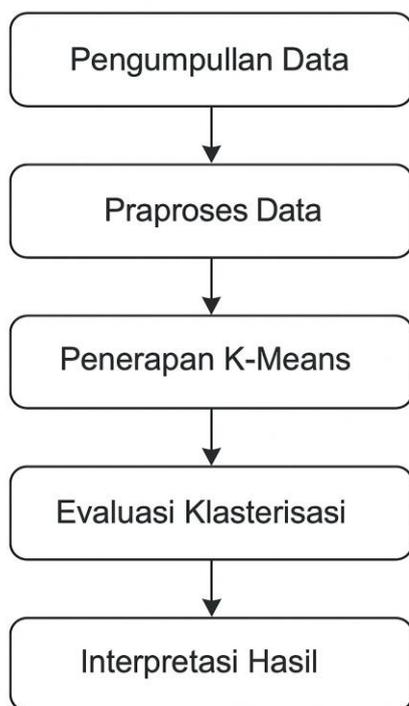
Penelitian sebelumnya yang berjudul "Solving Hajj and Umrah Challenges Using Information and Communication Technology: A Survey" membahas pemanfaatan teknologi informasi dan komunikasi (TIK) untuk mengatasi berbagai permasalahan dalam pelaksanaan ibadah haji dan umrah (Showail, 2022). Sedangkan penelitian ini yang berjudul "Algoritma K-Means untuk Klusterisasi Jumlah Jamaah Haji yang Diberangkatkan ke Tanah Suci Mekkah Menurut Kabupaten/Kota di Provinsi Sumatera Utara" dengan menerapkan metode K-Means diharapkan dapat memberikan kontribusi dalam upaya pengelompokan data jemaah haji berdasarkan wilayah administratif. Dan Pengelompokan ini bertujuan untuk membantu perencanaan dan pengambilan keputusan yang lebih efektif, sehingga dapat mendukung penyelesaian sebagian permasalahan dalam penyelenggaraan ibadah haji, seperti distribusi kuota, logistik dan pelayanan haji.

Dalam konteks penelitian ini, algoritma K-Means digunakan untuk mengklusterisasi jumlah jemaah haji yang diberangkatkan ke Tanah Suci Mekah berdasarkan kabupaten/kota di Provinsi Sumatera Utara. Penggunaan machine learning, khususnya metode klusterisasi, diharapkan mampu mengungkap distribusi wilayah dengan karakteristik jumlah jemaah yang berbeda-beda, seperti klaster jumlah jemaah tinggi, sedang, dan rendah. Informasi ini penting untuk membantu perencanaan pelayanan haji yang lebih efektif dan berbasis data.

Penerapan K-Means dalam analisis data sosial keagamaan membuktikan bahwa machine learning dapat digunakan tidak hanya untuk bidang industri atau teknologi, tetapi juga untuk mendukung pengambilan keputusan di sektor sosial (Djaka Permana et al., 2023). Data jumlah jemaah yang digunakan dalam penelitian ini diambil dari publikasi resmi Badan Pusat Statistik (BPS) Provinsi Sumatera Utara tahun 2024.

2. METODE

Diagram Alur Metode Penelitian



Gambar 1: Flowchart Metode Penelitian

Penelitian ini menggunakan pendekatan K-Means clustering untuk mengklusterkan

data jumlah jemaah haji berdasarkan kabupaten/kota di Provinsi Sumatera Utara. Langkah-langkah yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Pengumpulan Data: Data jumlah jemaah haji yang diberangkatkan dari masing-masing kabupaten/kota di Provinsi Sumatera Utara selama beberapa tahun terakhir dikumpulkan dari sumber resmi seperti Kementerian Agama atau instansi terkait.
2. Pra-pemrosesan Data: Data yang dikumpulkan kemudian dibersihkan dan diproses untuk memastikan bahwa tidak ada data yang hilang atau tidak valid.
3. Penentuan Jumlah Klaster: Sebelum menerapkan algoritma K-Means, dilakukan analisis untuk menentukan jumlah klaster yang optimal dengan menggunakan metode seperti Elbow Method atau Silhouette Score.
4. Penerapan Algoritma K-Means: Setelah jumlah klaster ditentukan, algoritma K-Means diterapkan untuk mengelompokkan kabupaten/kota berdasarkan jumlah jemaah haji yang diberangkatkan. Proses klusterisasi ini dilakukan menggunakan perangkat lunak statistik atau pemrograman seperti Python dengan library scikit-learn.
5. Evaluasi dan Interpretasi Hasil: Hasil klusterisasi kemudian dievaluasi untuk melihat apakah klaster yang terbentuk sesuai dengan pola yang diharapkan dan memberikan wawasan yang bermanfaat mengenai distribusi jemaah haji di Sumatera Utara.

A. Algoritma Machine Learning

Algoritma Machine Learning didefinisikan sebagai algoritma yang digunakan untuk melatih model, dalam Machine Learning dibagi menjadi tiga jenis yang berbeda, yaitu

- a) *Supervised Learning* (dalam dataset ini diberi label dan teknik Regresi dan Klasifikasi digunakan),
- b) *Unsupervised Learning* (dalam dataset ini tidak diberi label dan teknik yang digunakan seperti

Dimensionality reduction dan *Clustering*) dan

- c) *Reinforcement Learning* (algoritma di mana model belajar dari setiap tindakannya)(Heryadi & Wahyono, 2020)

B. Clustering Cluster

Clustering Cluster adalah sebuah grup yang memiliki kemiripan tertentu. *Clustering* adalah salah satu metode machine learning unsupervised untuk mengelompokkan objek-objek menjadi beberapa group atau cluster berdasarkan kemiripan dengan obyek yang lainnya(Ishak, 2022).

C. K-Means Clustering

Algoritma *K-Means* didefinisikan sebagai metode *Unsupervised Learning* yang memiliki proses berulang di mana dataset dikelompokkan ke dalam k jumlah *cluster* atau subkelompok yang tidak tumpang tindih yang telah ditentukan, membuat titik-titik dalam *cluster* sedekat mungkin ketika mencoba untuk menjaga *cluster* pada ruang yang berbeda itu mengalokasikan titik data ke *cluster* sehingga jumlah jarak kuadrat antara *cluster centroid* dan titik data berada pada titik data minimal, pada posisi ini sentroid dari *cluster* adalah nilai rata-rata dari titik data yang ada dalam *cluster*(Amalina et al., 2022). Secara umum tahapan algoritma *K-Means* sebagai berikut(Nugraha et al., 2022):

1. Tentukan k buah *cluster*
2. Pilih sejumlah k buah objek secara acak yang akan dijadikan sebagai titik *centroid cluster*
3. Tentukan k buah *centroid* (titik tengah)
4. Kelompokkan obyek ke centroid cluster terdekat berdasarkan Rumus *Euclidean distance* (Ong, n.d.):

$$d_{ik} = \sqrt{\sum_i^n (C_{ij} - C_{kj})^2}$$

Dimana:

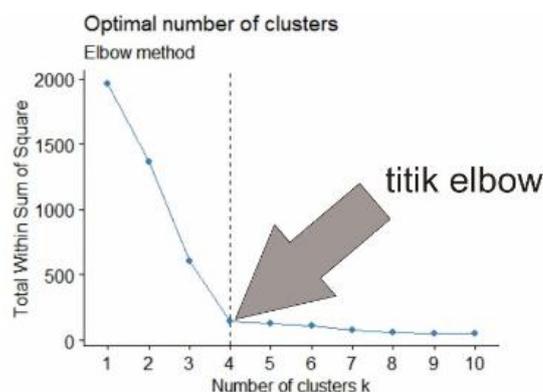
C_{ij} = pusat *cluster*

C_{kj} = data

5. Hitung Kembali semua titik *centroid*
6. Ulangi Langkah 3-5 sehingga nilai titik *centroid* tidak lagi berubah.

D. Metode elbow

Metode Elbow adalah salah satu metode atau teknik yang bisa diterapkan pada algoritma K-Means untuk menentukan jumlah K yang optimal(Maori, 2023) . Untuk menentukan jumlah K yang optimal dengan melihat titik pada graph dimana penurunan inersia tidak lagi signifikan seperti pada gambar 3 di bawah ini, jumlah K yang optimal adalah K=4.



Gambar 2. Penentuan jumlah *cluster* menggunakan *elbow*(Ayu et al., 2019).

3. HASIL DAN PEMBAHASAN

1. Data Awal

Berdasarkan data jumlah jemaah haji Badan Pusat Statistik Sumut 2024 yang diberangkatkan dari kabupaten/kota di Provinsi Sumatera Utara, Dataset berisi 33 kabupaten/kota di Sumatera Utara dengan data jumlah jemaah haji per wilayah dilakukan klasterisasi menggunakan algoritma K-Means untuk mengelompokkan kabupaten/kota dengan karakteristik serupa dalam hal jumlah jemaah haji.

Tabel 1. Tabel Data Jamah Haji (BPS) Sumut 2024

Kabupaten/Kota	Jemaah Haji
Nias	...
Mandailing Natal	760
Tapanuli Selatan	252
Tapanuli Tengah	112

Tapanuli Utara	8
Toba Samosir	3
Labuhan Batu	833
Asahan	400
Simalungun	169
Dairi	18
Karo	14
Deli Serdang	762
Langkat	470
Nias Selatan	...
Humbang Hasundutan	1
Pakpak Bharat	6
Samosir	...
Serdang Bedagai	319
Batu Bara	217
Padang Lawas Utara	167
Padang Lawas	263
Labuhan Batu Selatan	8
Labuhan Batu Utara	10
Nias Utara	2
Nias Barat	...
Kota Sibolga	52
Kota Tanjung Balai	184
Kota Pematang Siantar	191
Kota Tebing Tinggi	129
Kota Medan	2357
Kota Binjai	302
Kota Padangsidimpuan	493
Kota Gunungsitoli	14
Sumatera Utara	8516

Keterangan Data: (...) Data tidak tersedia

2. Praproses Data

Sebelum klasterisasi dilakukan, ada 3 data yang hilang (misalnya dari kabupaten/kota seperti Nias, Nias Selatan, dan Samosir).

3. Penerapan K-Means / (Klasterisasi)

Berikut adalah Tabel Perhitungan Klasterisasi K-Means berdasarkan jumlah jemaah haji yang diberangkatkan ke Tanah Suci dari kabupaten/kota di Provinsi Sumatera Utara. Tabel ini akan mencakup pengelompokan berdasarkan jumlah jemaah haji dan klaster yang dihasilkan sesuai dengan jumlah jemaah menggunakan Google Colab.

Tabel 2. Distribusi Klaster Berdasarkan Jumlah Jemaah Haji

	Kabupaten/Kota	Jemaah Haji	Klaster
0	Mandailing Natal	760	3
1	Tapanuli Selatan	252	2
2	Tapanuli Tengah	112	2
3	Tapanuli Utara	8	1
4	Toba Samosir	3	1
5	Labuhan Batu	833	3
6	Asahan	400	2
7	Simalungun	169	2
8	Dairi	18	1
9	Karo	14	1
10	Deli Serdang	762	3
11	Langkat	470	2
12	Humbang Hasundutan	1	1
13	Pakpak Bharat	6	1
14	Serdang Bedagai	319	2
15	Batu Bara	217	2
16	Padang Lawas Utara	167	2
17	Padang Lawas	263	2
18	Labuhan Batu Selatan	8	1
19	Labuhan Batu Utara	10	1
20	Nias Utara	2	1
21	Kota Sibolga	52	1
22	Kota Tanjung Balai	184	2
23	Kota Pematang Siantar	191	2
24	Kota Tebing Tinggi	129	2
25	Kota Medan	2357	3
26	Kota Binjai	302	2
27	Kota Padangsidimpuan	493	2
28	Kota Gunungsitoli	14	1
29	Sumatera Utara	8516	3

Berikut adalah Penjelasan Tabel 2 Distribusi klaster berdasarkan jumlah jemaah haji: Klaster

1. Klaster 1 (Jumlah Jemaah Haji Rendah): Kab/Kota dengan jumlah jemaah haji antara 0 hingga 100. Contoh: Tapanuli Utara (8), Toba Samosir (3), Pakpak Bharat (6), dan Humbang Hasundutan (1).
2. Klaster 2 (Jumlah Jemaah Haji Sedang): Kab/Kota dengan jumlah jemaah haji antara 100 hingga 500. Contoh: Tapanuli Selatan (252), Asahan (400), Kota Tebing Tinggi (129), dan Kota Padangsidimpuan (493).
3. Klaster 3 (Jumlah Jemaah Haji Tinggi): Kab/Kota dengan jumlah jemaah haji lebih dari 500. Contoh: Kota Medan (2357), Labuhan Batu (833), Deli Serdang (762), dan Sumatera Utara (Total) yang memiliki 8516 jemaah.

Setelah menentukan jumlah kluster yang optimal, didapatkan 3 kluster berdasarkan jumlah jemaah haji yang tercatat pada masing-masing kabupaten/kota. Penggunaan Elbow Method untuk menentukan jumlah kluster yang optimal mengindikasikan bahwa 3 kluster adalah jumlah yang paling sesuai.

4. Evaluasi Klusterisasi:

Lalu dilakukan proses perhitungan manual menggunakan rumus *Euclidean Distance*:

$$d_{ik} = \sqrt{\sum_i^n (C_{ij} - C_{kj})^2}$$

Menghitung manual Euclidean Distance menggunakan centroid awal berikut:

- **C1 (Rendah) = 50**
- **C2 (Sedang) = 300**
- **C3 (Tinggi) = 520**

1. Mandailing Natal (x = 760)

- Jarak Ke C1 = $\sqrt{(760) - (50)^2} = \sqrt{710^2} = 710$
- Jarak Ke C2 = $\sqrt{(760) - (300)^2} = \sqrt{460^2} = 460$
- Jarak Ke C3 = $\sqrt{(760) - (520)^2} = \sqrt{240^2} = 240$

Kluster: C3 (Tinggi), karena jarak ke C3 yang paling kecil. Dari ketiga sampel data dapat di lanjutkan perhitungan dengan menggunakan data jarak terdekat dengan Centroid. Dapat dilihat hasilnya pada Tabel berikut.

Tabel 3. Hasil Kluster menggunakan Google Colab

Kabupaten/Kota	Jemaah Haji	Jarak ke C1	Jarak ke C2	Jarak ke C3	Kluster
Mandailing Natal	760	710	460	240	3
Tapanuli Selatan	252	202	48	748	2
Tapanuli Tengah	112	62	188	888	1
Tapanuli Utara	8	42	292	992	1
Toba Samosir	3	47	297	997	1
Labuhan Batu	833	783	533	167	3
Asahan	400	350	100	600	2
Simalungun	169	119	131	831	1
Dairi	18	32	282	982	1
Karo	14	36	286	986	1
Deli Serdang	762	712	462	238	3
Langkat	470	420	170	530	2
Humbang Hasundutan	1	49	299	999	1
Pakpak Bharat	6	44	294	994	1
Serdang Bedagai	319	269	19	681	2
Batu Bara	217	167	83	783	2
Padang Lawas Utara	167	117	133	833	1
Padang Lawas	263	213	37	737	2
Labuhan Batu Selatan	8	42	292	992	1
Labuhan Batu Utara	10	40	290	990	1
Nias Utara	2	48	298	998	1
Kota Sibolga	52	2	248	948	1
Kota Tanjung Balai	184	134	116	816	2
Kota Pematang Siantar	191	141	109	809	2
Kota Tebing Tinggi	129	79	171	871	1
Kota Medan	2357	2307	2057	1357	3
Kota Binjai	302	252	2	698	2
Kota Padangsidimpuan	493	443	193	507	2
Kota Gunungsitoli	14	36	286	986	1
Sumatera Utara	8516	8466	8216	7516	3

Hitung Centroid Baru

Setelah semua daerah masuk kluster baru, hitung rata-rata setiap kluster.

- **Kluster 1 (Rendah): Anggota:** Tapanuli Utara (8), Toba Samosir (3), Dairi (18), Karo (14), Humbang Hasundutan (1), Pakpak Bharat (6), Labuhan Batu Selatan (8), Labuhan Batu Utara (10), Nias Utara (2), Kota Sibolga (52), Kota Gunungsitoli (14)
 Total: 8+3+18+14+1+6+8+10+2+52+14= 136
 Jumlah anggota: 11
 Centroid baru C1 = 136/11 ≈ 12,36
- **Kluster 2 (Sedang): Anggota:** Tapanuli Selatan (252), Tapanuli Tengah (112), Asahan (400), Simalungun (169), Langkat (470), Serdang Bedagai (319), Batu Bara (217), Padang Lawas Utara (167), Padang Lawas (263), Kota Tanjung Balai (184), Kota Pematang Siantar (191), Kota Tebing Tinggi (129), Kota Binjai (302), Kota

Padangsidempuan (493) Total: 3668
 Jumlah anggota: 14
 Centroid baru C2 = $3668/14 \approx 262$

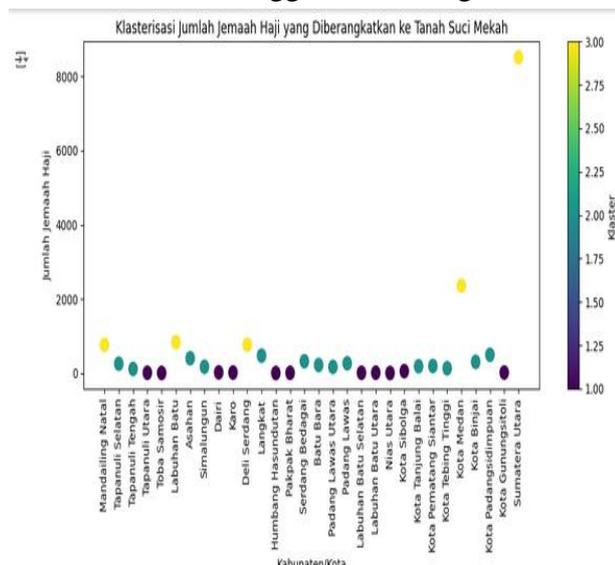
- Klaster 3 (Tinggi): Anggota: Mandailing Natal (760), Labuhan Batu (833), Deli Serdang (762), Kota Medan (2357), Sumatera Utara (8516) Total: 15228
 Jumlah anggota: 5
 Centroid baru C3 = $15228/5 \approx 3045,6$

5. Interpretasi Hasil:

- Klaster 1 = daerah dengan rata-rata **sekitar 12 jemaah** (sangat kecil)
- Klaster 2 = daerah dengan rata-rata **sekitar 262 jemaah** (sedang)
- Klaster 3 = daerah dengan rata-rata **sekitar 3045 jemaah** (sangat besar)

Jadi, pemisahan yang tadi sudah **sangat sesuai** dengan hasil manual **K-Means**.

K-Means merupakan algoritma klasterisasi berbasis jarak yang bertujuan untuk mengelompokkan data ke dalam k klaster berdasarkan kedekatan atau kemiripan terhadap pusat klaster (centroid). Berikut ini adalah hasil visualisasi proses klasterisasi yang telah dilakukan menggunakan Google Colab.



Gambar 3. Visualisasi Klasterisasi

4. PENUTUP

4.1. Kesimpulan

Penelitian ini menunjukkan bahwa algoritma K-Means dapat digunakan untuk mengelompokkan jumlah jemaah haji yang diberangkatkan dari kabupaten/kota di Provinsi Sumatera Utara. Hasil klasterisasi membagi daerah menjadi beberapa kelompok berdasarkan jumlah jemaah, seperti kelompok tinggi, sedang, dan rendah. Klaster ini membantu dalam memahami pola distribusi jemaah dan bisa digunakan sebagai dasar perencanaan dan pengambilan kebijakan oleh instansi terkait.

4.2. Saran

Penelitian selanjutnya disarankan untuk menambahkan data yang lebih informatif seperti usia dan jenis kelamin jemaah, serta menggabungkan metode K-Means dengan visualisasi peta agar hasil klaster lebih jelas dan bermanfaat bagi pengambil keputusan.

5. DAFTAR PUSTAKA

Amalina, T., Bima, D., Pramana, A., & Sari, B. N. (2022). Metode K-Means Clustering Dalam Pengelompokan Penjualan Produk Frozen Food. *Jurnal Ilmiah Wahana Pendidikan*, 8(15), 574–583. <https://doi.org/10.5281/zenodo.7052276>

Ayu, D., Dewi, I. C., & Pramita, K. (2019). Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali. In *JURNAL MATRIX* (Vol. 9, Issue 3).

Dakhi, Y. L., & Ningsi, B. A. (2024). Pengelompokan Kabupaten dan Kota Provinsi Sumatera Utara Berdasarkan Indikator Kesejahteraan Rakyat Menggunakan Algoritma K-Means. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(3), 993–1003. <https://doi.org/10.57152/malcom.v4i3.1381>

Djaka Permana, M., Lia Hananto, A., Novalia, E., Huda, B., & Paryono, T. (2023). Klasterisasi Data Jamaah Umrah pada

- Tanurmutmainah Tour Menggunakan Algoritma K-Means. *Jurnal KomtekInfo*, 15–20.
<https://doi.org/10.35134/komtekinfo.v10i1.332>
- Heryadi, Y., & Wahyono, T. (2020). *Machine Learning: Konsep dan Implementasi*.
<https://www.researchgate.net/publication/344419764>
- Hidayat, T. (2022). Klasifikasi Data Jamaah Umroh Menggunakan Metode K-Means Clustering. *Jurnal Sistim Informasi Dan Teknologi*, 19–24.
<https://doi.org/10.37034/jsisfotek.v4i1.115>
- Humaira Mufidah, T., Risky, M., Anshori, Z., & Jenderal Penyelenggaraan Haji dan Umrah Kementerian Agama, D. (2023). Ciledug Raya No.99, RT.10/RW.2, Petukangan Utara. *Jurnal Media Infotama*, 19(1), 3–4.
- Ishak, R. (2022). *Clustering Tingkat Pemahaman Dasar Mahasiswa Pada Pra-Perkuliahan Probabilitas Statistika Dengan Metode K-Means*. 4.
- Maori, N. A. (2023). METODE ELBOW DALAM OPTIMASI JUMLAH CLUSTER PADA K-MEANS CLUSTERING. *Jurnal SIMETRIS*, 14.
- Nasution, Mf., & Roesnelly, R. (2024). *CLASSIFICATION OF K-NEAREST NEIGHBOR (K-NN) AND CONVOLUTIONAL NEURAL NETWORK (CNN) FOR THE IDENTIFICATION OF BRONCHITIS DISEASE IN TODDLERS USING GLCM FEATURE EXTRACTION BASED ON THORAX X-RAY IMAGES* (Vol. 02, Issue 01).
- Nugraha, A., Nurdiawan, O., & Dwilestari, G. (2022). PENERAPAN DATA MINING METODE K-MEANS CLUSTERING UNTUK ANALISA PENJUALAN PADA TOKO YANA SPORT. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 6, Issue 2).
- Nusa, T., Komputasi, J., & Informasi, T. (n.d.). 36 *Sebagai Jurnal Terakreditasi Peringkat 4 berdasarkan Surat Keputusan Dirjen Risbang SK Nomor 85/M/KPT/2020*.
<https://doi.org/10.33480/techno.v22i1.6367>
- Ong, J. O. (n.d.). *IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING UNTUK MENENTUKAN STRATEGI MARKETING PRESIDENT UNIVERSITY*.
- Showail, A. J. (2022). Solving Hajj and Umrah Challenges Using Information and Communication Technology: A Survey. *IEEE Access*, 10, 75404–75427.
<https://doi.org/10.1109/ACCESS.2022.3190853>
- Wilayah, P., Di, D., Riau, P., Algoritma, M., Tata, K.-M., Pertiwi, A., Halim, F. R., Afni, N. ', Fahrezi, I. R., & Pramana, D. (n.d.). *SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat Mapping of Da'wah Areas in Riau Province Using the K-Means Algorithm*.
<https://journal.irpi.or.id/index.php/sentimas>
- Zhao, Y., & Zhou, X. (2021). K-means Clustering Algorithm and Its Improvement Research. *Journal of Physics: Conference Series*, 1873(1).
<https://doi.org/10.1088/1742-6596/1873/1/012074>