

PEMETAAN TEMPAT IBADAH DI JAKARTA: SEBUAH STUDI MENGUNAKAN KMEANS CLUSTERING

Intan Aulia Putri^{1*}, Fadiana², Daniel Yeri Kristiaanto³

^{1,2,3}Institut Teknologi Telkom Purwokerto

12311103145@ittelkom-pwt.ac.id¹, 2211103055@ittelkom-pwt.ac.id²

daniel@ittelkom-pwt.ac.id³

Received: 01-05-2024

Revised: 10-05-2024

Approved: 16-5-2024

ABSTRAK

Tempat ibadah adalah fasilitas umum untuk menunjang kegiatan beribadah setiap umat beragama agar setiap warga yang beribadah merasa nyaman. Indonesia mengakui agama Islam, Kristen, Katolik, Budha, Hindu dan Konghucu sebagai agama yang sah secara hukum. Jakarta merupakan ibukota negara Indonesia yang memiliki jumlah penduduk 11.23 juta jiwa pada tahun 202. Adanya keberagaman agama dan jumlah penduduk yang terus meningkat maka dilakukan pemetaan tempat ibadah untuk 6 agama di Jakarta. Penelitian ini dilakukan untuk mengetahui penyebaran tempat ibadah 6 agama di Jakarta dan mengetahui cara Metode K-Means dalam proses clustering penyebaran tempat ibadah. Hasil diperoleh bahwa jumlah tempat ibadah terbanyak ada di wilayah Jakarta Pusat dan Jakarta Selatan. Persebaran dengan jumlah sedang berada di Jakarta Timur, Jakarta Barat, dan Jakarta Utara serta persebaran sedikit berada di Kepulauan Seribu.

Kata Kunci: Clustering, Data Mining, K-Means.

PENDAHULUAN

Tempat ibadah adalah fasilitas umum yang ada di setiap daerah di Indonesia. Tempat ibadah biasanya dibangun dan disediakan oleh pemerintahan daerah di Indonesia. Fasilitas ini dibangun guna menunjang kegiatan beribadah setiap umat beragama. Seluruh umat beragama berharap setiap wilayah memiliki sarana dan prasarana yang memadai agar setiap warga yang beribadah merasa nyaman. Indonesia merupakan negara yang memiliki keberagaman agama. Hal tersebut merupakan salah satu kekayaan bangsa yang dijamin dalam Pancasila sila pertama "Ketuhanan Yang Maha Esa" dan pada UUD 1945 pasal 29 ayat 2. Bangsa Indonesia mengakui 6 agama resmi diantaranya Agama Islam, Katolik, Protestan, Hindu, Budha dan Konghucu.

DKI Jakarta memiliki beberapa kota administrasi. Jakarta terbagi menjadi lima wilayah kota madya dan satu kabupaten administratif diantaranya, Kota madya Jakarta Pusat, Jakarta Utara, Jakarta Barat, Jakarta Selatan, dan Jakarta Timur, serta Kabupaten Administratif Kepulauan Seribu. Menurut data Dukcapil tahun 2022, jumlah penduduk di Jakarta sebanyak 10.748.230 jiwa. Tahun 2023, jumlah penduduk di wilayah Jakarta meningkat sebesar 6.000 jiwa menjadi 11.337.563 jiwa. Di daerah Jakarta banyak pendatang dari berbagai wilayah di Indonesia. Dengan lonjakan jumlah penduduk setiap tahunnya maka adanya tempat ibadah yang memadai untuk setiap umat beragama harus diperhatikan. Adanya keberagaman agama dan jumlah penduduk yang terus meningkat maka dilakukan pemetaan tempat ibadah untuk enam agama di Jakarta.

Pemetaan tempat ibadah menjadi penting karena meningkatnya populasi dengan keberagaman agama. Metode K-Means Clustering akan digunakan untuk menganalisis dan memetakan persebaran tempat ibadah dari berbagai agama di Jakarta.

K-means Clustering menjadi metode yang efektif untuk pengelompokan data non-hierarki yang memisahkan data ke dalam bentuk cluster. Metode ini akan

mengelompokkan data yang memiliki karakteristik atau fitur yang sama dan memisahkan data dengan karakteristik yang berbeda (Amalina et al., 2022). Tujuan dari studi ini untuk memberi pemahaman lebih lanjut tentang pola distribusi tempat ibadah di Jakarta dan memberikan wawasan lebih serta pemahaman tentang keberagaman agama dalam perkotaan.

STUDI LITERATUR

Penelitian ini merujuk pada penelitian sebelumnya yang telah mengeksplorasi Clustering dengan menerapkan Algoritma K-Means. Sebagai contoh, penelitian oleh Sri yang berjudul "Pengklasteran Kabupaten/Kota di Jawa Tengah berdasarkan Tenaga Kesehatan dengan Menggunakan Metode Ward dan K-Means" menunjukkan pembentukan tiga kluster. Dalam penelitian tersebut, metode Ward menghasilkan nilai rasio simpangan baku sebesar 0,3019%, sedangkan metode K-Means menghasilkan nilai sebesar 0,2974%. Hasil menunjukkan bahwa metode K-Means lebih efektif daripada metode Ward [4]. Di sisi lain, penelitian yang dilakukan oleh Desi dengan judul "Clustering Jumlah Tenaga Kesehatan Berdasarkan Kecamatan di Kabupaten Karawang Menggunakan Algoritma K-Means" menerapkan algoritma K-Means Clustering menggunakan Microsoft Excel dan menghasilkan tiga cluster dengan nilai DBI sebesar -0.370 (Sitinjak et al., 2022). Selain itu, penelitian oleh Yuni yang berjudul "Implementasi Data Mining Dalam Mengelompokkan Jumlah Penduduk Miskin Berdasarkan Provinsi Menggunakan Algoritma K-Means" menemukan bahwa iterasi clustering pada data jumlah penduduk miskin dilakukan sebanyak tiga kali (Yuni Radana Sembiring et al., 2021).

Data Mining

Data Mining merupakan proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode. *Data mining* juga dapat disebut sebagai proses yang digunakan untuk menentukan sebuah struktur data (Suryanata et al., 2019).

Clustering

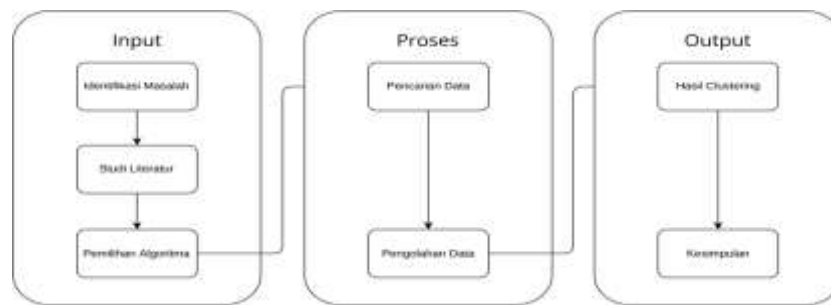
Clustering merupakan proses membagi data dalam suatu himpunan ke dalam beberapa kelompok yang kesamaan datanya dalam suatu kelompok lebih besar daripada kesamaan data tersebut dengan data dalam kelompok lain (Sadewo et al., 2018). Analisis *cluster* merupakan pengelompokan objek-objek data hanya berdasarkan pada informasi yang terdapat pada data, yang menjelaskan objek dan relasinya (Javed Mehedi Shamrat et al., 2020)

Algoritma K-Means

Algoritma K-Means adalah teknik pengelompokan data yang melibatkan iterasi berulang. Ini dapat mengelompokkan data dalam jumlah besar secara cepat dan efisien (Nishom, 2019). K-Means merupakan salah satu contoh algoritma clustering yang termasuk dalam kategori pembelajaran tanpa pengawasan (Nabila et al., 2021) KMeans membagi data menjadi beberapa kelompok, mengelompokkan data serupa bersama, dan merupakan algoritma sederhana yang populer dan mudah dimengerti (Adha et al., 2021)

METODE PENELITIAN

Pada penelitian ini bertujuan untuk melakukan pengelompokan pada data persebaran tempat ibadah di Kota Administrasi Jakarta Utara pada tahun 2018 menggunakan algoritma K-Means, dengan menggunakan Microsoft Excel.



Gambar 1. Alur Penelitian

Identifikasi Masalah

Pada langkah ini, penulis mengidentifikasi masalah yang ada untuk menetapkan tema penelitian. Identifikasi masalah merupakan langkah awal dalam proses penelitian. Hal ini mencakup upaya untuk menemukan masalah serta menentukan strategi yang dapat dilakukan untuk mengatasinya.

METODE PENELITIAN

Pada tahap ini dilakukan metode penelitian bertujuan untuk mencari sumber referensi, yang berguna untuk penentuan metode atau algoritma, mencari informasi dari sumber penelitian terdahulu. Metode yang digunakan adalah K-Means dalam proses clustering penyebaran tempat ibadah. Untuk rumus perhitungan sebagai berikut:

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj}$$

Keterangan:

- V_{ij} = Centroid rata-rata cluster ke-i untuk variable ke-j
- N_i = Jumlah data yang menjadi anggota cluster ke-i
- i, k = Indeks dari cluster
- j = Indeks dari variabel
- X_{kj} = Nilai data ke-k variable ke-j untuk cluster tersebut

Euclidean Distance

Euclidean distance, yang digunakan secara luas untuk mengevaluasi kesamaan antara dua objek, telah terbukti lebih efektif daripada metode lain seperti Manhattan dan Minkowski distance (Hasan, 2015) [10].

$$De = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$$

Keterangan:

- De = Eucclidean Distance
- i = Banyaknya objek
- (x,y) = Koordinat objek
- (s,t) = Koordinat centroid

Pemilihan Algoritma

Proses pemilihan algoritma melibatkan pemilihan algoritma yang akan digunakan dalam penelitian ini. Dalam studi ini, algoritma yang dipilih adalah K-Means, sebuah algoritma data mining yang digunakan untuk tujuan pengelompokan.

Pencarian Data

Pada langkah ini, dilakukan identifikasi data yang akan digunakan dalam

penelitian ini. Proses pencarian data bertujuan untuk memperoleh dataset yang relevan untuk penelitian. Sumber data yang digunakan berasal dari Jakarta.bps.go.id, dengan fokus pada data Jumlah Sarana Ibadah Menurut Kabupaten/Kota Provinsi DKI Jakarta dari tahun 2020 hingga 2022.

Pengolahan Data Clustering

Pada tahap ini, data yang telah diperoleh akan diolah menggunakan perangkat lunak Microsoft Excel. Berikut adalah langkah-langkah urutan dalam melakukan proses pengelompokan data menggunakan Microsoft Excel:

1. Tentukan k sebagai jumlah cluster yang akan dibentuk, dengan menggunakan metode kriteria siku (elbow criterion) yang dihitung menggunakan rumus berikut:

$$SSE = \sum_{k=1}^K \sum_{x_i \in S_k} \|N_i - C_k\|^2$$

2. Pilih titik pusat cluster (centroid) awal sebanyak k, yang dipilih secara acak dari objek yang tersedia. Penentuan awal centroid dilakukan dengan cara acak dari objek yang ada sebanyak k cluster. Untuk menghitung centroid cluster berikutnya, digunakan rumus berikut:

$$v = \frac{\sum_{i=1}^n X_i}{n} : i = 1, 2, 3, \dots, n$$

3. Perhitungan jarak dari setiap objek ke centroid masing-masing cluster dilakukan menggunakan Euclidean Distance, dengan rumus sebagai berikut:

$$De = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$$

4. Objek-objek dialokasikan ke centroid terdekat masing-masing. Dalam iterasi umum, alokasi objek ke setiap cluster dilakukan dengan metode hard k-means, di mana setiap objek secara tegas ditetapkan sebagai anggota cluster berdasarkan jaraknya ke pusat cluster.
5. Dilakukan iterasi untuk menentukan posisi centroid baru dengan menggunakan persamaan.
6. Langkah ketiga diulang jika posisi centroid baru tidak sama.

Hasil Clustering

Pada tahap ini dihasilkan sebuah data hasil dari proses clustering data mining menggunakan algoritma *k-means*.

HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini mencakup Jumlah Sarana Tempat Ibadah dari tahun 2020 hingga 2022 di Provinsi DKI Jakarta. Atribut yang tersedia dalam data ini meliputi Kabupaten/Kota, Masjid, Musholla, Gereja Protestan, Gereja Katolik, Pura, Vihara, dan Kelenteng. Pengujian akan dilakukan pada enam wilayah yang berbeda.

Tabel 1.
Data jumlah sarana tempat ibadah di Provinsi DKI Jakarta

No.	Kota	Masjid	Musholla	Gereja Protestan	Gereja Katolik	Pura	Vihara	Kelenteng
1	Kep Seribu	36	99	0	0	0	0	0
2	Jakarta Selatan	1926	742	513	18	13	23	1
3	Jakarta Timur	2506	2579	909	35	15	53	2

No.	Kota	Masjid	Musholla	Gereja Protestan	Gereja Katholik	Pura	Vihara	Kelenteng
4	Jakarta Pusat	1563	1374	643	20	22	73	2
5	Jakarta Barat	2250	2525	748	39	4	295	2
6	Jakarta Utara	1715	1953	672	27	20	200	2

Normalisasi Data

Normalisasi data bertujuan untuk mengurangi redundansi data dan memastikan data terstruktur dengan benar dalam tabel. Data yang telah diperoleh akan dinormalisasi menggunakan rumus z-score di Excel. Di bawah ini terdapat rumus z-score dan tabel data yang telah dinormalisasi:

$$Z = (x - \mu) / \sigma$$

Keterangan :

x = nilai yang diamati (skor mentah)

μ = rata-rata populasi

σ = adalah standar deviasi populasi

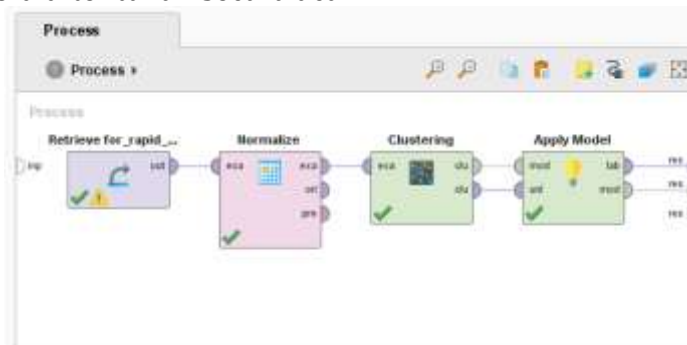
Z = Z Score (Nilai Baku)

Tabel 1.
Data Normalisasi

Kota	Masjid	Musholla	Gereja Prote...	Gereja Kath...	Pura	Vihara	Kelenteng
Kep Seribu	-1.874	-1.452	-1.856	-1.656	-1.412	-0.930	-1.793
Jakarta Selat...	0.299	-0.806	-0.217	-0.369	0.076	-0.731	-0.598
Jakarta Timur	0.966	1.038	1.048	0.846	0.305	-0.471	0.598
Jakarta Pusat	-0.118	-0.172	0.199	-0.226	1.107	-0.298	0.598
Jakarta Barat	0.671	0.984	0.534	1.132	-0.954	1.626	0.598
Jakarta Utara	0.056	0.408	0.291	0.274	0.878	0.803	0.598

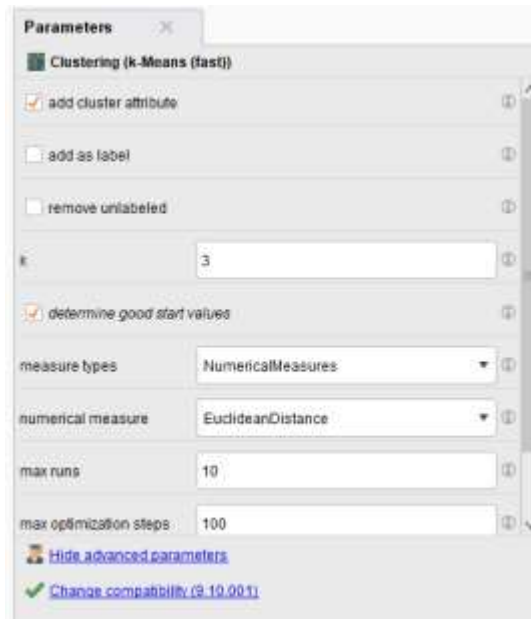
Proses Clustering menggunakan Rapid Miner

Setelah semua data dinormalisasi, data dapat dikelompokkan menggunakan teknik k-means clustering. Tahap awal dari metode K-Means Clustering adalah menetapkan jumlah cluster, dalam kasus ini, sebanyak 3 cluster. Kemudian, pusat cluster atau centroid ditentukan secara acak.



Gambar 2. Model Rapid Miner

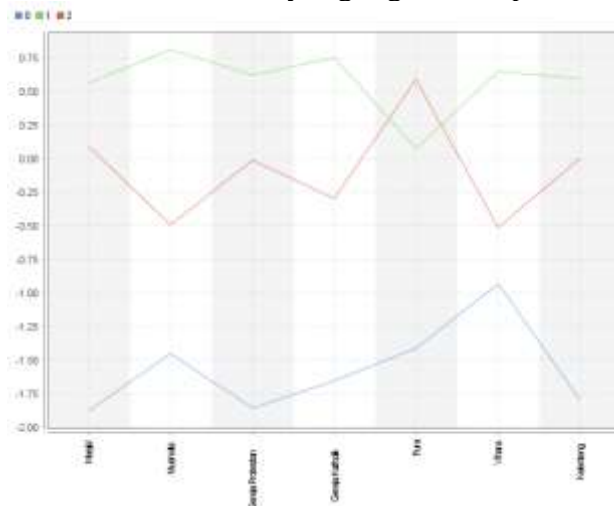
Tetapkan nilai parameter yang ada pada bagian clustering. Isi parameter k = 3, measure types = Numerical Measures dan numerical measure = Euclidean Distance. Data yang diproses akan menghasilkan 3 model cluster.



Gambar 4. Set Parameter K-Means

Attribute	cluster_0	cluster_1	cluster_2
Masjid	-1.874	0.564	0.090
Musholla	-1.452	0.810	-0.489
Gereja Prote...	-1.856	0.625	-0.009
Gereja Kath...	-1.656	0.750	-0.298
Pura	-1.412	0.076	0.592
Vihara	-0.930	0.653	-0.514
Kelenteng	-1.793	0.598	0

Gambar 5. data centroid yang digunakan pada rapid miner.



Gambar 6. Hasil pemetaan menggunakan rapid miner

Tabel.
 Hasil cluster data tempat ibadah enam agama di Indonesia

id	cluster	Kota	Masjid	Musholla	Gereja Prote...	Gereja Kath...	Para	Vihara	Kelenteng
1	cluster_0	Kep Seribu	-1.874	-1.452	-1.856	-1.556	-1.412	-0.930	-1.793
2	cluster_2	Jakarta Selat...	0.299	-0.806	-0.217	-0.369	0.076	-0.731	-0.598
3	cluster_1	Jakarta Timur	0.966	1.038	1.049	0.846	0.305	-0.471	0.598
4	cluster_2	Jakarta Pusat	-0.118	-0.172	0.199	-0.226	1.107	-0.298	0.598
5	cluster_1	Jakarta Barat	0.671	0.984	0.534	1.132	-0.954	1.626	0.598
6	cluster_1	Jakarta Utara	0.056	0.409	0.291	0.274	0.878	0.803	0.598

Setelah proses pengolahan data, ditemukan bahwa terdapat 1 wilayah dalam cluster 0, 3 wilayah dalam cluster 1, dan 2 wilayah dalam cluster 2. Dari hasil tersebut, terungkap bahwa Jakarta Pusat dan Jakarta Selatan menonjol sebagai kota dengan jumlah tempat ibadah terbanyak (cluster 2), diikuti oleh Jakarta Timur, Jakarta Barat, dan Jakarta Utara yang menjadi wilayah dengan jumlah tempat ibadah terbanyak kedua (cluster 1). Wilayah terakhir, yaitu Kepulauan Seribu, termasuk dalam cluster 0 sebagai wilayah dengan jumlah tempat ibadah terkecil.

KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, maka proses clustering (pengelompokan) penyebaran tempat ibadah enam Agama di Provinsi DKI Jakarta tahun 2020 -2022 didapatkan hasil bahwa jumlah persebaran terbanyak berada di wilayah Jakarta Pusat dan Jakarta Selatan. Persebaran dengan jumlah sedang berada di Jakarta Timur, Jakarta Barat, dan Jakarta Utara serta persebaran sedikit berada di Kepulauan Seribu.

DAFTAR PUSTAKA

- Adha, R., Nurhaliza, N., Sholeha, U., & Mustakim, M. (2021). Perbandingan Algoritma DBSCAN dan K-Means Clustering untuk Pengelompokan Kasus Covid-19 di Dunia. *SITEKIN: Jurnal Sains, Teknologi Dan Industri*, 18(2), 206–211. <https://ejournal.uin-suska.ac.id/index.php/sitekin/article/view/12469>
- 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.
- Javed Mehedi Shamrat, F. M., Tasnim, Z., Mahmud, I., Jahan, N., & Nobel, N. I. (2020). Application of k-means clustering algorithm to determine the density of demand of different kinds of jobs. *International Journal of Scientific and Technology Research*, 9(2), 2550–2557.
- Nabila, Z., Isnain, A. R., Permata, & Abidin, Z. (2021). Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means. *Jurnal Teknologi Dan Sistem Informasi (JTSI)*, 2(2), 100. <http://jim.teknokrat.ac.id/index.php/JTSI>
- Nishom, M. (2019). Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square. *Jurnal Informatika: Jurnal Pengembangan IT*, 4(1), 20–24. <https://doi.org/10.30591/jpit.v4i1.1253>
- Sadewo, M. G., Eriza, A., Windarto, A. P., Hartama, D., & Prodi. (2018). Algoritma K-Means Dalam Mengelompokkan Desa / Kelurahan Menurut Keberadaan Keluarga

- Pengguna Listrik dan Sumber Penerangan Jalan Utama Berdasarkan Provinsi. *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS) SAINTEKS 2019*, 754–761.
- Sitinjak, D. K., Pangestu, B. A., & Sari, B. N. (2022). Clustering Tenaga Kesehatan Berdasarkan Kecamatan di Kabupaten Karawang Menggunakan Algoritma K-Means. *Journal of Applied Informatics and Computing*, 6(1), 47–54. <https://doi.org/10.30871/jaic.v6i1.3855>
- Suryanata, M. G., Pane, D. H., & Hutasuhut, M. (2019). Implementasi Algoritma K-Means Dalam Mengukur Tingkat Kepuasan Siswa Terhadap Pelayanan Sekolah Studi Kasus Di Mts.S Nurul Amaliyah. *Jurnal Teknologi Sistem Informasi Dan Sistem Komputer TGD*, 2(2), 118–125.
- Yuni Radana Sembiring, Saifullah, & Riki Winanjaya. (2021). Implementasi Data Mining Dalam Mengelompokkan Jumlah Penduduk Miskin Berdasarkan Provinsi Menggunakan Algoritma. *KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen) Vol. 2, No. 2, 2(2)*, 125–132.