E-ISSN : 2988-1986
https://ejournal.cahayailmubangsa.institute/index.php/kohesi

Kohesi: Jurnal Multidisiplin Saintek
Volume 11 No 2 Tahun 2026

PENERAPAN METODE CLUSTERING DALAM MENGELOMPOKKAN PERMOHONAN SURAT KETERANGAN DI KUA (STUDI KASUS: KUA BINJAI SELATAN)

Dwi Winda Salsabila [1*]
[1,2,3] Program Studi Sistem Informasi, STMIK Kaputama
E-mail: dwiwindasalsabila26@gmail.com

ABSTRACT

The South Binjai Religious Affairs Office (KUA) receives an increasing number of certificate applications every year, but the data has only been stored as archives and has not been used to support decision-making. This study aims to apply a clustering method to group certificate application patterns based on applicant age, sub-district of residence, and type of letter submitted. The method used is the K-Means algorithm through the Knowledge Discovery in Database (KDD) stages, namely data selection, preprocessing, transformation, data mining, and evaluation, assisted by the MATLAB application. The optimal number of clusters was determined using the Elbow method based on the Sum of Squared Error (SSE) value, which produced an elbow point at k = 3. Certificate application data for the 2023-2025 period were then grouped into three clusters using the K-Means algorithm. The results show that the clustering process successfully formed three groups with distinct characteristics: Cluster 1 (589 records) dominated by applicants aged 29-36 in Binjai Estate sub-district applying for marriage-status certificates, while Cluster 2 (268 records) and Cluster 3 (308 records) were dominated by applicants aged 20-28 applying for marriage-intent letters. These clustering results can provide information for the South Binjai KUA to improve the effectiveness and efficiency of marriage administration services based on real data patterns.

Keywords: *Clustering, K-Means, Elbow Method, Certificate, KUA, Data Mining*

ABSTRAK

Kantor Urusan Agama (KUA) Binjai Selatan menerima permohonan surat keterangan yang jumlahnya terus meningkat setiap tahun, namun data permohonan tersebut selama ini hanya disimpan sebagai arsip dan belum dimanfaatkan untuk mendukung pengambilan keputusan. Penelitian ini bertujuan untuk

Article History
Received: Juli 2026
Reviewed:Juli 2026
Published: Juli 2026

Plagiarism Checker No 237.u873247.fh8374

Copyright : Author
Publish by : Kohesi

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

menerapkan metode clustering guna mengelompokkan pola permohonan surat keterangan berdasarkan variabel usia pemohon, kelurahan domisili, dan jenis surat yang diajukan. Metode yang digunakan adalah algoritma K-Means melalui tahapan Knowledge Discovery in Database (KDD), yaitu data selection, preprocessing, transformation, data mining, dan evaluation, dengan bantuan aplikasi MATLAB. Penentuan jumlah cluster optimal dilakukan menggunakan metode Elbow berdasarkan nilai Sum of Squared Error (SSE), yang menghasilkan titik siku pada k=3. Data permohonan surat keterangan periode 2023-2025 kemudian dikelompokkan ke dalam tiga cluster menggunakan algoritma K-Means. Hasil penelitian menunjukkan bahwa proses clustering berhasil membentuk tiga kelompok permohonan dengan karakteristik berbeda berdasarkan usia pemohon, kelurahan, dan jenis surat, dengan Cluster 1 (589 data) didominasi pemohon usia 29-36 tahun di Kelurahan Binjai Estate untuk surat Keterangan Sudah Menikah, sedangkan Cluster 2 (268 data) dan Cluster 3 (308 data) didominasi pemohon usia 20-28 tahun untuk surat Kehendak Menikah. Hasil pengelompokan ini dapat menjadi bahan informasi bagi KUA Binjai Selatan dalam meningkatkan efektivitas dan efisiensi layanan administrasi pernikahan berdasarkan pola data yang nyata.

Kata Kunci: *Clustering, K-Means, Metode Elbow, Surat Keterangan, KUA, Data Mining*

1. PENDAHULUAN

Kantor Urusan Agama (KUA) adalah unit di bawah Kementerian Agama Republik Indonesia yang bertugas memberikan pelayanan administrasi keagamaan Islam, terutama terkait pencatatan dan urusan pernikahan. Salah satu layanan rutin yang ditawarkan adalah penerbitan berbagai surat keterangan, seperti surat rekomendasi nikah, surat kehendak menikah, dan surat keterangan telah menikah. Permohonan surat-surat ini datang dari masyarakat dengan latar belakang yang beragam, mulai dari usia pemohon, daerah tempat tinggal, hingga jenis surat yang diminta.

Seiring meningkatnya kebutuhan masyarakat akan layanan administrasi pernikahan, jumlah permohonan surat keterangan di KUA Binjai Selatan terus bertambah setiap tahun. Sayangnya, data permohonan ini selama ini hanya disimpan sebagai arsip biasa dan belum dianalisis lebih dalam untuk menemukan pola yang dapat mendukung pengambilan keputusan. Penerapan teknik analisis data seperti clustering pada pelayanan publik dapat membantu mengevaluasi data historis dan menghasilkan informasi yang berguna bagi instansi (Hariyanto et al., 2025). Jika digali lebih lanjut, data ini dapat memberikan wawasan berharga, misalnya jenis surat yang paling sering diminta, kelompok usia yang paling aktif, dan daerah dengan permohonan tertinggi (Islam et al., 2025).

Clustering merupakan salah satu teknik dalam data mining yang termasuk kategori unsupervised learning, yaitu mengelompokkan data tanpa memerlukan label kelas sebelumnya, berdasarkan tingkat kemiripan karakteristik antarobjek (Wani, 2024; Jelita et al., 2023). Salah satu algoritma clustering yang paling banyak digunakan adalah K-Means, karena konsepnya sederhana, efisien, dan mudah diimplementasikan pada berbagai jenis data. Algoritma ini bekerja dengan membagi data ke dalam beberapa cluster berdasarkan jarak terdekat ke pusat cluster (centroid), dan telah banyak digunakan dalam penelitian analisis data administrasi, pendidikan, maupun layanan masyarakat (Febriyani et al., 2025).

Berdasarkan uraian tersebut, penelitian ini menerapkan metode clustering dengan algoritma K-Means untuk mengelompokkan data permohonan surat keterangan di KUA Binjai Selatan berdasarkan tiga variabel, yaitu usia pemohon, kelurahan domisili, dan jenis surat yang diajukan, dengan tujuan menghasilkan informasi pola permohonan yang dapat mendukung peningkatan kualitas dan efisiensi layanan administrasi pernikahan.

Rumusan masalah dalam penelitian ini adalah: (1) bagaimana karakteristik data permohonan surat keterangan di KUA Binjai Selatan dilihat dari usia pemohon, kelurahan, dan jenis surat; (2) bagaimana cara menerapkan metode clustering untuk mengelompokkan pola permohonan tersebut; dan (3) apa hasil pengelompokan yang diperoleh dari penerapan metode clustering. Penelitian ini bertujuan untuk menerapkan algoritma K-Means Clustering pada data permohonan surat keterangan KUA Binjai Selatan tahun 2023-2025 serta menentukan jumlah cluster optimal menggunakan metode Elbow, sehingga dihasilkan pengelompokan pola permohonan yang dapat dijadikan bahan pertimbangan bagi pihak KUA.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan data mining dengan teknik clustering melalui tahapan Knowledge Discovery in Database (KDD), yang meliputi lima tahap, yaitu data selection, preprocessing, transformation, data mining, dan evaluation/interpretation (Nuraeni Septiani & Saeful Anwar, 2023).

2.1 Data Selection

Data yang digunakan dalam penelitian ini adalah data Surat Permohonan KUA Binjai Selatan periode tahun 2023-2025 dengan jumlah 1000 data. Data tersebut memiliki atribut nama pemohon, tanggal permohonan, usia pemohon, kelurahan, dan jenis surat. Dari keseluruhan atribut tersebut, dipilih tiga atribut yang digunakan sebagai variabel penelitian, yaitu usia pemohon, kelurahan, dan jenis surat.

2.2 Data Preprocessing

Tahap preprocessing dilakukan untuk membersihkan data dari kesalahan penulisan, duplikasi, dan data yang tidak lengkap, sehingga data memiliki format yang konsisten dan siap

digunakan pada tahap berikutnya. Dari 1000 data awal, data yang telah melalui proses seleksi dan pembersihan digunakan sebagai dasar transformasi data pada tahap selanjutnya.

2.3 Transformation

Pada tahap ini, data kategorikal diubah menjadi bentuk numerik agar dapat diproses menggunakan algoritma K-Means. Variabel usia pemohon dikelompokkan ke dalam lima rentang dengan kode 1-5 (20-28, 29-36, 37-44, 45-52, dan ≥53 tahun), variabel kelurahan dikodekan ke dalam kode 1-8 sesuai delapan kelurahan wilayah kerja KUA Binjai Selatan (Bhakti Karya, Binjai Estate, Pujidadi, Rambung Barat, Rambung Dalam, Rambung Timur, Tanah Merah, dan Tanah Seribu), dan variabel jenis surat dikodekan menjadi tiga kategori, yaitu Rekomendasi (1), Keterangan Sudah Menikah (2), dan Kehendak Menikah (3).

Usia Pemohon	Kode	Kelurahan	Kode	Jenis Surat	Kode
20-28	1	Bhakti Karya	1	Rekomendasi	1
29-36	2	Binjai Estate	2	Keterangan Sudah Menikah	2
37-44	3	Pujidadi	3	Kehendak Menikah	3
45-52	4	Rambung Barat	4		
≥ 53	5	Rambung Dalam	5		
		Rambung Timur	6		
		Tanah Merah	7		
		Tanah Seribu	8		

Tabel 1. Kode transformasi variabel usia, kelurahan, dan jenis surat

2.4 Data Mining (K-Means Clustering)

Algoritma K-Means mengelompokkan data melalui langkah-langkah berikut:

1. menentukan jumlah cluster (k);
2. menentukan centroid awal secara acak;
3. menghitung jarak setiap data ke masing-masing centroid menggunakan rumus Euclidean Distance;
4. mengelompokkan data ke cluster dengan jarak terdekat;
5. memperbarui nilai centroid berdasarkan rata-rata anggota cluster; dan

6. mengulangi proses hingga tidak terjadi perubahan anggota cluster (konvergen).

Rumus perhitungan centroid dan jarak Euclidean yang digunakan adalah sebagai berikut:

$$C(i)=\frac{X1+X2+X3+...+Xn}{\Sigma X} \quad ....(1)$$

$$Dij=\sqrt{(X1i-X1j)^2+(X2i-X2j)^2+...(Xki-Xkj)^2} \quad ....(2)$$

Pada penelitian ini, jumlah cluster ditetapkan sebanyak 3 cluster berdasarkan hasil metode Elbow, dengan centroid awal dipilih secara acak dari data alternatif, yaitu C1 = (4,8,1), yang selanjutnya diperbarui secara iteratif hingga konvergen. Proses perhitungan dilakukan dengan bantuan aplikasi MATLAB menggunakan fungsi kmeans dengan parameter jarak sqeuclidean dan lima kali replikasi (Replicates = 5) untuk memperoleh hasil pengelompokan yang stabil.

2.5 *Metode Elbow*

Metode Elbow digunakan untuk menentukan jumlah cluster (k) yang paling optimal dengan menghitung nilai Sum of Squared Error (SSE) pada beberapa nilai k, kemudian memplotkan nilai tersebut ke dalam grafik. Titik pada grafik yang membentuk siku (elbow), yaitu titik di mana penurunan nilai SSE mulai melandai, dipilih sebagai jumlah cluster optimal. Pada penelitian ini, titik siku terbentuk secara jelas pada k=3, sehingga nilai tersebut digunakan sebagai jumlah cluster dalam proses K-Means Clustering

2.6 *Evaluation*

Tahap evaluasi dilakukan untuk menginterpretasikan hasil pengelompokan data ke dalam informasi yang bermakna, meliputi analisis jumlah anggota tiap cluster, nilai centroid akhir tiap cluster, serta karakteristik dominan usia pemohon, kelurahan, dan jenis surat pada masing-masing cluster.

3. HASIL DAN PEMBAHASAN

3.Penentuan Cluster Optimal dengan Metode Elbow

Proses analisis diawali dengan menentukan jumlah cluster optimal menggunakan metode Elbow berdasarkan nilai SSE pada MATLAB. Hasil plot SSE terhadap beberapa nilai k menunjukkan penurunan nilai SSE yang signifikan hingga k=3, kemudian penurunan mulai melandai pada nilai k berikutnya. Titik siku (elbow point) yang terbentuk pada k=3 ini mengindikasikan bahwa tiga cluster merupakan jumlah pengelompokan yang paling optimal untuk data permohonan surat keterangan KUA Binjai Selatan.

3.2 Proses Clustering dengan K-Means

Data yang telah ditransformasi diimpor ke command window MATLAB, kemudian diproses menggunakan fungsi kmeans dengan k=3. Proses iterasi dilakukan hingga tidak terjadi perubahan anggota cluster antara satu iterasi dengan iterasi berikutnya (konvergen). Hasil visualisasi berupa grafik tiga dimensi menampilkan sebaran titik data pada masing-masing cluster berdasarkan tiga variabel, yaitu usia (sumbu X), kelurahan (sumbu Y), dan jenis surat (sumbu Z), beserta posisi centroid akhir tiap cluster.

3.3 Hasil Pengelompokan

Hasil pengujian dengan k=3 menghasilkan jumlah anggota masing-masing cluster serta nilai centroid akhir sebagai berikut:

Cluster	Jumlah Data	Centroid(Usia,Kelurahan, Jenis Surat)	Karakteristik Dominan
C1	589	1,58 ; 2,14 ; 2,46	Usia 29-36 th, Kel. Binjai Estate, Keterangan Sudah Menikah
C2	268	1,50 ; 4,89 ; 2,55	Usia 20-28 th, Kel. Rambung Dalam, Kehendak Menikah
C3	308	1,50 ; 7,56 ; 2,52	Usia 20-28 th, Kel. Tanah Seribu, Kehendak Menikah

Tabel 2. Hasil pengelompokan cluster, nilai centroid, dan karakteristik dominan

Cluster 1 memiliki jumlah anggota terbanyak, yaitu 589 data, dengan karakteristik dominan pemohon berusia 29-36 tahun yang berdomisili di Kelurahan Binjai Estate dan mengajukan surat Keterangan Sudah Menikah. Cluster ini merepresentasikan kelompok permohonan yang paling dominan di KUA Binjai Selatan. Cluster 2 beranggotakan 268 data dengan karakteristik dominan pemohon berusia 20-28 tahun di Kelurahan Rambung Dalam yang mengajukan surat Kehendak Menikah. Cluster 3 beranggotakan 308 data dengan karakteristik dominan pemohon berusia 20-28 tahun di Kelurahan Tanah Seribu, juga mengajukan surat Kehendak Menikah.

KESIMPULAN

Berdasarkan hasil analisis clustering terhadap data permohonan surat keterangan di KUA Binjai Selatan menggunakan metode K-Means dengan variabel usia pemohon, kelurahan, dan jenis surat, dapat disimpulkan bahwa: (1) data permohonan surat keterangan di KUA Binjai Selatan memiliki karakteristik yang beragam ditinjau dari usia pemohon, kelurahan asal, dan jenis surat yang dimohonkan, dengan variasi permohonan yang cukup signifikan antarkelurahan, kelompok usia, maupun jenis surat; (2) penerapan metode clustering dengan algoritma K-Means dilakukan melalui tahapan KDD (data selection, preprocessing, transformation, data mining, dan evaluation) dengan bantuan aplikasi MATLAB, di mana jumlah cluster optimal ditentukan

melalui metode Elbow berdasarkan nilai SSE dan diperoleh k=3; dan (3) hasil pengelompokan menghasilkan tiga cluster dengan karakteristik berbeda, yaitu Cluster 1 (589 data, usia 29-36 tahun, Kelurahan Binjai Estate, surat Keterangan Sudah Menikah) sebagai kelompok permohonan paling dominan, Cluster 2 (268 data) dan Cluster 3 (308 data) yang keduanya didominasi pemohon usia 20-28 tahun dengan jenis surat Kehendak Menikah pada kelurahan yang berbeda. Hasil pengelompokan ini dapat memberikan rekomendasi bagi KUA Binjai Selatan untuk meningkatkan efektivitas layanan administrasi pernikahan melalui pendekatan yang lebih tepat sasaran berdasarkan karakteristik masing-masing kelompok permohonan.

Untuk penelitian selanjutnya, disarankan untuk menggunakan metode lain seperti Naive Bayes atau memperluas variabel penelitian, misalnya waktu permohonan, jenis kelamin pemohon, dan status perkawinan, sehingga hasil pengelompokan yang diperoleh dapat lebih komprehensif dan representatif. Selain itu, disarankan agar pihak KUA Binjai Selatan dapat memanfaatkan hasil pengelompokan ini sebagai dasar penyusunan kebijakan dan perencanaan layanan administrasi pernikahan yang lebih terstruktur dan berbasis data.

DAFTAR PUSTAKA

[1] Aldi Febrian, Neny Marlina, S.IP., M. ., & Departemen. (2025). Pengaruh Pemberian Insentif Terhadap Kualitas Pelayanan Publik Pada Dinas Kependudukan dan Pencatatan Sipil Kota Jakarta Selatan.

[2] Aliyah, H. (2022). Implementasi Pelaksanaan Standar Pelayanan Minimal (SPM) Di Kantor Urusan Agama Kecamatan Mojoroto Kementrian Agama Kota Kediri. 11(1), 182-198.

[3] Candra, Y., Tri, A., & Dani, R. (2025). Transformasi Data Kementerian: Studi Clustering Hierarki pada Layanan Keagamaan. 7, 2075-2086.

[4] Dicha Mutia Dhani., Relita Buaton., I. G. P. (2024). Penerapan Metode K-Means Clustering untuk Menentukan Kepuasan Mahasiswa terhadap Fasilitas Sarana dan Prasarana Kampus di STMIK Kaputama Binjai. 3.

[5] Febriyani, S., Buroiroh, H., Wahyu, N., Septiani, N. T., & Faiz, M. (2025). Clustering Daerah Rawan Gempa di Pulau Jawa Berbasis Metode Algoritma K-Means. 5(2), 115-122.

[6] Hariyanto, Y., Primadewi, A., & Hanafi, M. (2025). Analisis Kepuasan Masyarakat terhadap Pelayanan Publik Menggunakan K-Means Clustering. 6(2). https://doi.org/10.47065/josh.v6i2.6577

[7] Islam, Y. I., Maulana, N. A., Azise, N., Ibrahimy, U., Timur, S. J., Ibrahimy, U., Timur, S. J., Ibrahimy, U., & Timur, S. J. (2025). Sistem Informasi Pembuatan Surat Rekom Pindah Nikah Berbasis Website. 2(1), 76-84.

[8] Jelita, T., Buaton, R., & Simajuntak, M. (2023). Pengelompokan Bidang Usaha Terhadap Bantuan Produktif Usaha Mikro (BPUM) Berdasarkan Wilayah Deli Serdang Menggunakan Metode Clustering K-Means (Studi Kasus: Dinas Koperasi Dan UMKM Kabupaten Deli Serdang). 3(2), 50-57.

[9] Magister, P., Keluarga, H., Iain, P., & Raya, P. (n.d.). Kebijakan Kementrian Agama Dalam Memberikan Layanan Nikah Bagi Daerah 3T Di Kabupaten Pulang Pisau. 1586-1607.

[10] Meiliana, A., Hutasoit, B., Buaton, R., & Sitepu, K. A. B. (2025). Data Mining Faktor Penyebab Masuk Penjara Menggunakan Metode Clustering (Studi Kasus: Lembaga Pemasyarakatan Kelas 1 Medan). 1, 662-672.

[11] Nuraeni Septiani, Saeful Anwar, R. H. (2023). Penerapan Algoritma K-Means Clustering Untuk Harga Rumah di Jakarta Selatan. 1(2).

[12] Rahmawati, N., Buaton, R., Ambarita, I., Jl, A., No, V., Kota, K. B., Binjai, K., & Utara, S. (2024). Klasifikasi Tingkat Kepuasan Masyarakat Penggunaan BPJS Kesehatan di Kota Binjai Menggunakan K-Means Clustering. 5.

[13] Renni, A., & Ambarita, Novrianini, I. (2023). The Grouping Of Types Of Tax Revenue In Binjai City Uses The K-Means Clustering Method Algorithm. 3(1).

[14] Riani, A. P., Voutama, A., & Ridwan, T. (2023). Penerapan K-Means Clustering Dalam Pengelompokan Hasil Belajar Peserta Didik Dengan Metode Elbow. 6, 164-172.

[15] Sulistiyawati, A., & Supriyanto, E. (2021). Implementasi Algoritma K-means Clustring dalam Penetuan Siswa Kelas Unggulan. 15(2), 25-36.

[11] Wani, A. A. (2024). Comprehensive Analysis Of Clustering Algorithms: Exploring Limitation and Innovative Solution. https://doi.org/10.7717/peerj-cs.2286