

# ALGORITMA K-MEANS DALAM PENGELOMPOKAN SURAT KELUAR DI KANTOR KEMENTERIAN AGAMA KOTA PEMATANG SIANTAR

<sup>1</sup>Mhd Aditia Nasution\*, <sup>2</sup>M. Safii

<sup>1,2</sup>Sistem Informasi, Stikom Tunas Bangsa Pematang Siantar  
Jln. Surya, Pematang Siantar, Sumatera Utara, Indonesia

\*e-mail: [mhdaditians2901@gmail.com](mailto:mhdaditians2901@gmail.com)<sup>1</sup>, [m.safii@amiktunasbangsa.ac.id](mailto:m.safii@amiktunasbangsa.ac.id)<sup>2</sup>

## Abstrak

Pengelolaan surat keluar di lingkungan Kantor Kementerian Agama Kota Pematang Siantar merupakan aspek krusial dalam menjalankan operasionalnya. Dalam upaya untuk meningkatkan efisiensi dan efektivitas pengelolaan surat keluar, penelitian ini memperkenalkan implementasi Algoritma K-Means sebagai salah satu solusi. Tujuan penelitian ini adalah untuk mengevaluasi dampak penggunaan Algoritma K-Means dalam pengelompokan surat keluar. Hasil penelitian menunjukkan bahwa Algoritma K-Means telah membantu secara signifikan dalam proses pengelompokan surat keluar dengan mengidentifikasi pola dan karakteristik yang relevan. Pengelompokan surat keluar berdasarkan kesamaan atribut tertentu telah memungkinkan peningkatan efisiensi distribusi surat, serta memberikan pemahaman yang lebih baik tentang pola pengiriman. Selain itu, potensi pengembangan lebih lanjut dalam hal integrasi dengan sistem manajemen surat elektronik sudah terlihat. Dengan demikian, hasil penelitian ini membuktikan bahwa Algoritma K-Means memiliki potensi untuk meningkatkan pengelolaan surat keluar di lingkungan Kantor Kementerian Agama Kota Pematang Siantar dan dapat dianggap sebagai alat yang berharga dalam meningkatkan efisiensi operasional dan pelayanan publik.

**Kata kunci:** Pengelolaan surat, Pengelompokan surat, Algoritma K-Means, Data Mining, Clustering.

## Abstract

*The management of outgoing letters within the Office of the Ministry of Religious Affairs in Pematang Siantar City is a crucial aspect of its operations. In an effort to enhance the efficiency and effectiveness of outgoing letter management, this research introduces the implementation of the K-Means Algorithm as one of the solutions. The objective of this research is to evaluate the impact of using the K-Means Algorithm in grouping outgoing letters. The results of the research indicate that the K-Means Algorithm has significantly aided in the process of grouping outgoing letters by identifying relevant patterns and characteristics. Grouping outgoing letters based on specific attribute similarities has allowed for improved letter distribution efficiency and a better understanding of sending patterns. Furthermore, the potential for further development in terms of integration with an electronic letter management system is evident. Therefore, this research demonstrates that the K-Means Algorithm has the potential to enhance outgoing letter management within the Office of the Ministry of Religious Affairs in Pematang Siantar City and can be considered a valuable tool in improving operational efficiency and public service.*

**Keywords:** Letter management, Letter Grouping, K-Means Algorithm, Data Mining, Clustering.



DOI: <https://doi.org/10.52362/jmijayakarta.v4i1.1304>

## 1 Pendahuluan (or Introduction)

Algoritma K-Means merupakan salah satu metode clustering yang dapat digunakan untuk mengelompokkan suatu kumpulan data berdasarkan karakteristik tertentu yang ada pada kumpulan data tersebut. Konsep K-Means mengelompokkan data dengan cara membagi data menjadi dua kelompok atau lebih.[1]

Surat keluar adalah salah satu bentuk komunikasi tertulis yang telah digunakan secara luas dalam berbagai aspek kehidupan, termasuk dalam konteks bisnis, administrasi pemerintahan, pendidikan, dan komunikasi pribadi. Surat keluar adalah dokumen tertulis yang dikirim oleh satu pihak kepada pihak lain dengan tujuan menyampaikan informasi, pesan, atau komunikasi tertentu[2]. Penggunaan surat keluar telah menjadi bagian integral dari komunikasi manusia selama berabad-abad, dan dengan kemajuan teknologi, bentuknya telah berkembang dari surat fisik yang dikirim melalui pos menjadi surat elektronik yang dapat dikirim secara instan melalui email.

Pengelompokan surat keluar memudahkan dalam pengarsipan dan pengelompokan sesuai dengan keperluan yang ada pada organisasi[3][4]. Dari 2 cluster yang digunakan dalam penerapan algoritma K-Means didapat bahwa metode ini dapat mengelompokkan surat keluar. Penelitian ini dilakukan untuk mengetahui penggunaan algoritma K-Means terhadap sekumpulan data data surat keluar pada Kantor Kementerian Agama yang dibagi menjadi 2 kelompok yakni kepegawaian (KP) dan keuangan (KU).

Pengelolaan surat keluar merupakan salah satu aspek yang sangat penting dalam operasional setiap instansi, termasuk Kantor Kementerian Agama Kota Pematang Siantar. Surat keluar mencakup berbagai dokumen, komunikasi, dan informasi yang harus diatur, diproses, dan didistribusikan dengan efisien dan efektif. Dalam konteks ini, penggunaan teknologi dan algoritma data mining menjadi semakin relevan untuk membantu meningkatkan pengelolaan surat keluar.

## 2 Tinjauan Literatur (or Literature Review)

### 2.1 Eksplorasi Algoritma K-Means:

Membahas secara mendalam tentang konsep dasar dari algoritma K-Means. Menyajikan literatur terkait penggunaan K-Means dalam bidang pengelompokan data.

### 2.2 Penerapan Algoritma K-Means pada Pengelompokan Surat Keluar:

Menyelidiki bagaimana algoritma K-Means dapat diterapkan secara efektif dalam konteks pengelompokan surat keluar di Kantor Kementerian Agama Kota Pematang Siantar. Menguraikan langkah-langkah konkret implementasi algoritma K-Means pada data surat keluar.

### 2.3 Keunggulan Algoritma K-Means:

Meneliti keunggulan algoritma K-Means dalam hal efisiensi dan efektivitas dibandingkan dengan metode pengelompokan data lainnya. Membahas dampak positif penerapan K-Means dalam meningkatkan keteraturan dan kecepatan pengelompokan surat keluar.

### 2.4 Analisis Hasil Pengelompokan:

Melakukan analisis mendalam terhadap hasil pengelompokan surat keluar menggunakan algoritma K-Means. Mengidentifikasi pola-pola atau hubungan antar kelompok surat keluar yang muncul setelah penerapan algoritma K-Means.

### 2.5 Evaluasi Kinerja Algoritma:

Melakukan evaluasi kinerja algoritma K-Means, baik dari segi waktu eksekusi maupun akurasi pengelompokan. Membandingkan hasil evaluasi dengan metrik-metrik kinerja yang umumnya digunakan dalam penelitian sejenis.



DOI: <https://doi.org/10.52362/jmijayakarta.v4i1.1304>

## 2.6 Implikasi Praktis:

Mendiskusikan implikasi praktis dari penggunaan algoritma K-Means dalam pengelompokan surat keluar di Kantor Kementerian Agama. Memberikan rekomendasi untuk pengembangan atau penyempurnaan lebih lanjut dalam implementasi algoritma.

## 2.7 Perspektif Pengguna:

Menyajikan pandangan dan tanggapan dari pengguna atau pemangku kepentingan terkait penggunaan algoritma K-Means dalam pengelompokan surat keluar. Menyelidiki sejauh mana algoritma ini memenuhi kebutuhan dan harapan dari pihak terkait.

## 3 Metode Penelitian (or Research Method)

Ada banyak metode yang bisa digunakan dalam clustering, salah satu contohnya adalah metode K-Means. Pengelompokan dapat digunakan sebagai pengelompokan non-hierarki untuk membagi data menjadi dua kelompok atau lebih. K-Means merupakan suatu metode analisis kelompok yang hasilnya membagi N objek yang diamati menjadi K kelompok (cluster) dan masing-masing objek yang diamati termasuk dalam kelompok yang mempunyai nilai mean (median) terdekat[5].

Dalam clustering diperoleh hasil berupa keanggotaan setiap data pada setiap cluster dan data dominan untuk masuk ke dalam cluster tertentu. Algoritma K-Means untuk clustering data adalah sebagai berikut:

1. Tentukan jumlah kelompok.
2. Alokasikan data ke dalam kelompok secara acak.
3. Hitung pusat kelompok (rata-rata) dari data yang terdapat pada masing-masing kelompok.

Lokasi titik pusat pada setiap kelompok yang diambil dari rata-rata (*mean*) semua nilai data pada setiap fiturnya harus dihitung kembali. Untuk menghitung sentroid fitur ke-*i* dapat digunakan seperti pada persamaan 1.

$$C_i = \frac{1}{M} \sum_{i=1}^M x_i \quad \dots\dots\dots (1)$$

Keterangan :

C = Centroid

M = jumlah data dalam sebuah kelompok

*i* = fitur ke-*i* dalam sebuah kelompok  $x =$   
matrik set data M x N.

N = jumlah fitur persamaan 1 dilakukan sebanyak *p* (dimensi data ) dimensi sehingga *i* mulai 1 sampai *p*.



DOI: <https://doi.org/10.52362/jmijayakarta.v4i1.1304>

4. Alokasikan masing-masing data ke sentroid/rata-rata terdekat.
5. Kembali ke langkah 3, apabila masih ada data yang berpindah kelompok atau apabila ada perubahan nilai sentroid di atas nilai ambang yang ditentukan, atau apabila perubahan nilai pada fungsi objektif yang digunakan masih di atas nilai ambang yang ditentukan.

Dalam mengukur jarak pada ruang jarak (*distance space*) dengan cara Euclidean yaitu menggunakan formula seperti pada persamaan 2.

$$D(x_2, x_1) = ||x_2 - x_1 ||_2 = \sqrt{\sum_{i=1}^p |x_{2i} - x_{1i}|^2} \quad \dots\dots\dots (2)$$

Keterangan :

D = Jarak antara data  $x_2$  dan  $x_1$

$| \cdot |$  = Nilai mutlak

p = dimensi data

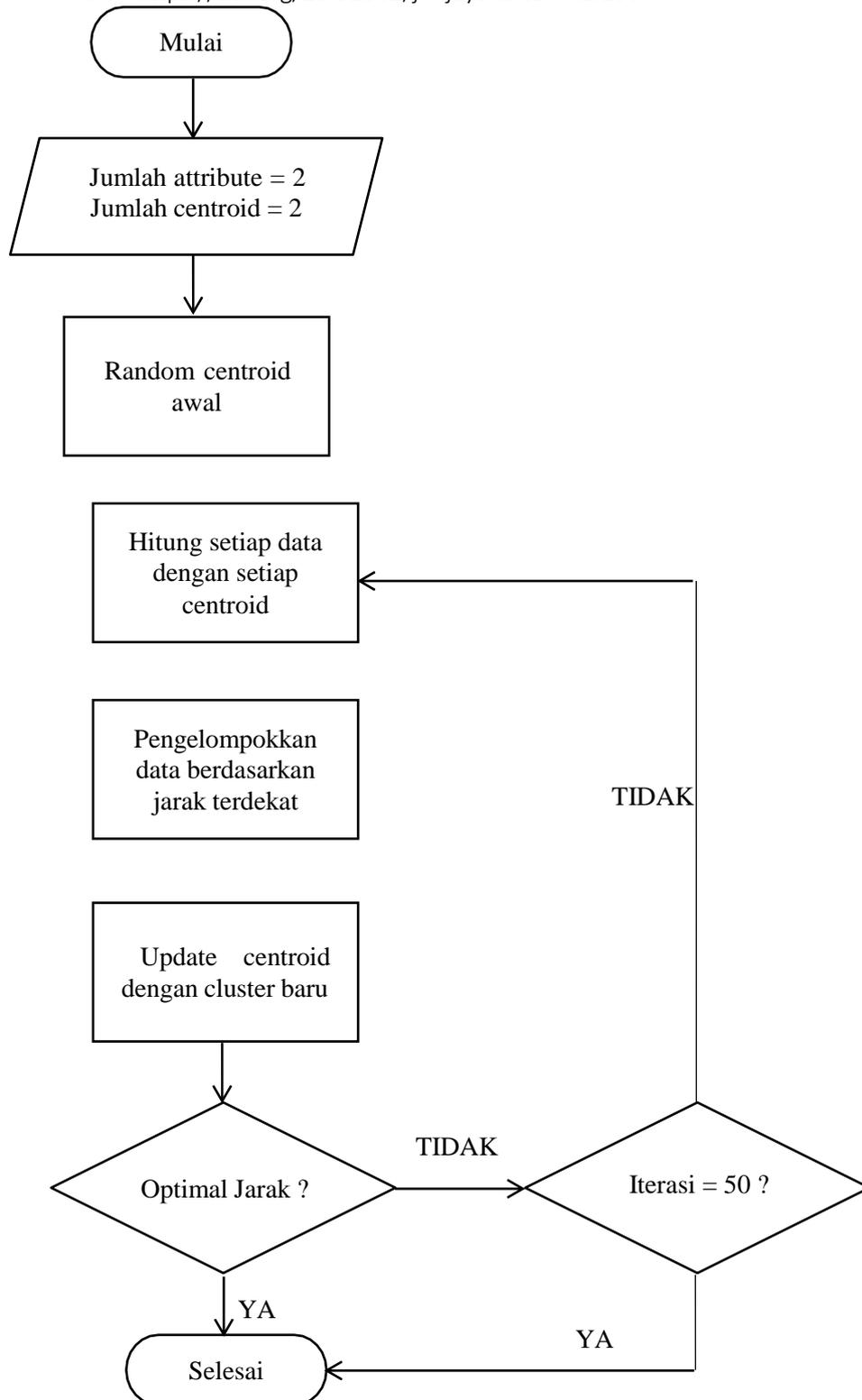
i = fitur ke-i dalam sebuah kelompok x = matrik set data M x N.

N = jumlah fitur

Gambar 1 menunjukkan *flowchart* konstruksi K-Means. Yakni alur proses dari konstruksi algoritma K-Means.



DOI: <https://doi.org/10.52362/jmijayakarta.v4i1.1304>



**Gambar 1** Flowchart Konstruksi K-Means



DOI: <https://doi.org/10.52362/jmijayakarta.v4i1.1304>

### 3.1 *Pengelolaan Surat Keluar*

Pengelolaan surat keluar adalah proses penting dalam administrasi kantor yang melibatkan pengiriman dokumen, surat, atau komunikasi kepada pihak yang bersangkutan. Pengelolaan surat keluar melibatkan beberapa tahap, termasuk:

1. Pengumpulan Informasi: Informasi tentang surat keluar, seperti jenis surat, tujuan, prioritas, dan tanggal pengiriman, harus dikumpulkan dengan baik.
2. Pengelompokan: Surat keluar dapat dikelompokkan berdasarkan berbagai kriteria, seperti tujuan pengiriman, jenis surat, atau prioritas. Pengelompokan ini dapat memudahkan proses distribusi.
3. Distribusi: Surat keluar harus didistribusikan kepada pihak yang bersangkutan sesuai dengan prioritas dan waktu yang ditentukan.
4. Pemantauan: Pemantauan proses pengiriman dan penerimaan surat keluar penting untuk memastikan keberhasilan pengelolaan surat keluar

### 3.2 *Microsoft Excel*

*Microsoft Excel* dapat digunakan untuk mengolah data dan memiliki fitur analisis. *Microsoft Excel* juga dapat digunakan untuk menyusun model data.

Dalam penelitian kali ini, *Microsoft Excel* sangat berperan dalam menjalankan algoritma K-Means yakni mulai dari menyusun model data, mengolah data, hingga menganalisis hasil dari model yang digunakan, yakni K-Means. *Microsoft Excel* memberikan kemudahan untuk perhitungan manual berdasarkan pada tahapan dari algoritma K-Means.

Beberapa fitur *Microsoft Excel* yang digunakan dalam penelitian ini, antara lain seperti berikut ini :

1. Sheet

Sheet pada *Microsoft Excel* diperlukan untuk membedakan masing-masing tahapan seperti tahapan inisialisasi dan iterasi yang berulang.

2. Sum

Perintah Sum digunakan untuk menambahkan data pada beberapa kolom atau baris.

3. Sumif

Perintah Sumif digunakan untuk menambahkan data sesuai kriteria tertentu.

4. SQRT

Perintah SQRT digunakan Untuk menghitung akar kuadrat.

## 4 Hasil dan Pembahasan (or Results and Analysis)

Pengelompokan menggunakan algoritma *K-Means* melibatkan 210 data yang telah melewati preprocessing dari data awal sebanyak 300 data surat keluar. Untuk menghindari data yang tidak lengkap dan yang tidak digunakan terkait penelitian ini. Pengelompokan 210 data pada set data pengelompokan 2 dimensi[6]. Dimensi data dari 3 fitur yaitu nomor surat (berdasarkan keperluan surat yaitu KP atau KU), status penerima surat, perihal surat. Nilai masing-masing data yang digunakan seperti pada tabel 1. Pengukuran jarak yang digunakan yaitu jarak Euclidean. Jumlah kelompok adalah 2 yaitu kepegawaian (KP) dan keuangan (KU). Ambang batas atau threshold (T) yang digunakan untuk perubahan fungsi objektif adalah 0,1.



DOI: <https://doi.org/10.52362/jmijayakarta.v4i1.1304>

**Tabel 1 Contoh nilai yang digunakan untuk pengelompokan**

Data ke-i	No.Surat	Status Penerima Surat	Perihal Surat
1	2	2	1
2	3	1	2
3	63	1	2

Pada tabel 1, untuk data perihal surat berdasarkan sepuluh bagian yaitu bobot surat dari skala 1 sampai 10 berdasarkan seringnya diperlukan surat tersebut. Untuk data status penerima surat berdasarkan tiga bagian yaitu pemberian bobot berdasarkan seringnya diperlukan surat tersebut dari skala 1 hingga 3.

**a. Inisialisasi**

- 1) Mengalokasikan semua data pada satu kelompok secara acak dari proses pengacakan yang dilakukan maka masing-masing data bergabung ke masing-masing kelompok dari 2 kelompok yakni kepegawaian (KP) dan keuangan (KU).
- 2) Menghitung pusat kelompok (sentroid). Lokasi pusat kelompok hasil dari persamaan 1 untuk menghitung pusat kelompok dari masing-masing fitur yang diambil dari rata-rata semua data yang bergabung dalam setiap kelompok

**Tabel 2 Menghitung pusat kelompok pada kelompok 1**

Cluster 1	Total Data	No. Surat	Status Penerima Surat	Perihal Surat
Jumlah	61	19831	166	390
Rata-rata		149,12	1,08	3,53

**Tabel 3 Menghitung pusat kelompok pada kelompok 2**

Cluster 2	Total Data	No. Surat	Status Penerima Surat	Perihal Surat
Jumlah	149	11064	178	541
Rata-rata		147,59	1,05	3,18

**Tabel 4 Pusat kelompok pada langkah inisiasi**

Kelompok	No. Surat	Status Penerima Surat	Perihal Surat
1	149,12	1,08	3,53
2	147,59	1,05	3,18



DOI: <https://doi.org/10.52362/jmijayakarta.v4i1.1304>

Nilai rata-rata pada tabel 2 dan tabel 3 digunakan pada tabel 4 dalam mencari pusat kelompok dan hasilnya untuk menghitung jarak data ke pusat kelompok seperti persamaan 2 dan didapat fungsi objektif yaitu sebesar 16997,78 sehingga perubahan fungsi objektif sebagai berikut :

$$\begin{aligned} \text{Jarak lama} &= 0 \\ \text{Jarak baru} &= 16997,78 \end{aligned}$$

$$\begin{aligned} \text{Perubahan fungsi objektif} &= |\text{jarak baru} - \text{jarak lama}| \\ &= |16997,78 - 0| \\ &= 16997,78 \end{aligned}$$

Nilai perubahan fungsi objek sebesar 17397,78 masih di atas nilai ambang batas (threshold) yaitu 0,1 maka pusat kelompok dihitung kembali ke iterasi selanjutnya sampai nilai perubahan fungsi objektif di bawah ambangbatas atau tidak adanya perpindahan kelompok.

b. Iterasi 1

- 1) Menghitung kembali jarak terpendek setiap data pada kelompok tertentu.
- 2) Mengalokasikan setiap data pada sentroid terdekat
- 3) Menghitung pusat kelompok (sentroid)

Seperti pada proses inisialisasi didapat pusat kelompok untuk iterasi 1 dan nilai rata-rata untuk masing-masing kelompok seperti pada tabel 5 dan tabel 6 Pusat kelompok pada iterasi 1 seperti pada tabel 7.

**Tabel 5 Menghitung pusat kelompok pada kelompok 1**

Cluster 1	Total Data	No. Surat	Status Penerima Surat	Perihal Surat
Jumlah	71	15831	158	483
Rata-rata		220,13	1,09	3,33

**Tabel 6 Menghitung pusat kelompok pada kelompok 2**

Cluster 2	Total Data	No. Surat	Status Penerima Surat	Perihal Surat
Jumlah	139	15064	143	460
Rata-rata		73,17	1,03	3,31



**Tabel 7 Pusat kelompok pada iterasi 1**

Kelompok	No. Surat	Status Penerima Surat	Perihal Surat
1	220,13	1,09	3,33
2	73,17	1,03	3,31

Nilai rata-rata digunakan pada tabel 8 dan didapat fungsi objektif yaitu sebesar 25477,52 sehingga perubahan fungsi objektif sebagai berikut:

$$\begin{aligned} \text{Jarak lama} &= 16997,78 \\ \text{Jarak baru} &= 25477,52 \\ \text{Perubahan fungsi objektif} &= |\text{jarak baru} - \text{jarak lama}| \\ &= |25477,52 - 16997,78| \\ &= 8479,74 \end{aligned}$$

Lanjut ke iterasi berikutnya dengan langsung menampilkan langkah menghitung pusat kelompok. Untuk mendapatkan jarak baru pada tahapan iterasi ke-1.

**c. Iterasi 2**

Pusat kelompok untuk iterasi 2 seperti pada tabel 9 dan tabel 10 dan nilai rata-rata untuk kelompok 1 seperti pada tabel 11.

**Tabel 1 Menghitung pusat kelompok pada kelompok 1**

Cluster 1	Total Data	No. Surat	Status Penerima Surat	Perihal Surat
Jumlah	71	15831	158	483
Rata-rata		220,13	1,09	3,33

**Tabel 2 Menghitung pusat kelompok pada kelompok 2**

Cluster 2	Total Data	No. Surat	Status Penerima Surat	Perihal Surat
Jumlah	139	15064	143	460
Rata-rata		73,17	1,03	3,31



**Tabel 3 Pusat kelompok pada iterasi 2**

Kelompok	No. Surat	Status Penerima Surat	Perihal Surat
1	220,13	1,09	3,33
2	73,17	1,03	3,31

Fungsi objektif yaitu sebesar 26477,52 sehingga perubahan fungsi objektif sebagai berikut:

$$\text{Jarak lama} = 25477,52$$

$$\text{Jarak baru} = 25477,52$$

$$\begin{aligned} \text{Perubahan fungsi objektif} &= |\text{jarak baru} - \text{jarak lama}| \\ &= |25477,52 - 25477,52| \\ &= 0,00 \end{aligned}$$

Pada iterasi ke-2, nilai perubahan fungsi objek sebesar 0,00 berada di bawah nilai ambang batas (threshold) yaitu 0,1 atau tidak adanya perpindahan kelompok maka iterasi dihentikan.

## 5 Kesimpulan (or Conclusion)

Kesimpulan yang dapat diambil dari penggunaan algoritma K-Means pada kumpulan data surat keluar yang terkluster adalah algoritma K-Means dapat digunakan untuk mengelompokkan surat keluar berdasarkan jumlah surat, status penerima pesan dan subjek surat dengan menerapkan langkah dan ketentuan algoritma K-Means, khususnya dengan menggunakan iterasi. Iterasi terus dilakukan selama nilai perubahan fungsi benda tetap lebih besar dari nilai ambang batas yang diberikan, yaitu 0,1, atau gerak kelompok masih terjadi. Pada iterasi kedua dapat dihentikan karena nilainya kurang dari nilai ambang batas yaitu 0,00 dan tidak ada lagi pergerakan kelompok.

## Referensi (Reference)

- [1] T. F. S. Aji, Mugiarto, and K. F. Ramdhania, "Sistem Informasi Pengarsipan Surat Masuk dan Keluar dengan Algoritma Sequential Search di Kelurahan Bahagia," *J. Inform. Inf. Secur.*, vol. 3, no. 1, pp. 111–124, 2022, doi: 10.31599/jiforty.v3i1.1336.
- [2] N. Riyana and F. Fatmawati, "Penerapan Algoritma Naive Bayes Untuk Klasifikasi Penerima Bantuan Surat Keterangan Tidak Mampu," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 2, no. 4, pp. 206–213, 2021, doi: 10.35746/jtim.v2i4.117.
- [3] M. F. Sjojfan and R. T. Aldisa, "Kombinasi Algoritma Sequential Search dan Fitur Autocomplete Pada Aplikasi Arsip (E-Arsip) Perpustakaan Berbasis Web," *J. Inf. Syst. Res.*,



DOI: <https://doi.org/10.52362/jmijayakarta.v4i1.1304>

vol. 4, no. 4, pp. 1262–1269, 2023, doi: 10.47065/josh.v4i4.3748.

- [4] I. Pratiwi, “Implementasi Algoritma Left Corner Parsing Dalam Pencarian Surat Pada Arsip Surat Masuk dan Surat Keluar,” vol. 1, no. 3, pp. 99–102, 2022.
- [5] Y. Muharam and Y. Bahtiar, “Pembangunan Aplikasi Klasifikasi Kode Surat Berbasis Android Menggunakan Algoritma Boyer-Moore Di Kantor Kecamatan Ciparay,” *J. Inform.*, vol. 08, no. 4, pp. 50–57, 2021.
- [6] I. M. Rian Yulianan, D. Care Khrisne, and P. Arya Mertasana, “Penerapan Algoritma K-Means Clustering dalam Penentuan Nilai Huruf pada Permainan Susun Kata Bahasa Bali,” *J. SPEKTRUM*, vol. 6, no. 3, p. 87, 2019, doi: 10.24843/spektrum.2019.v06.i03.p12.

