JITET (Jurnal Informatika dan Teknik Elektro Terapan)
Vol. 13 No. 1, pISSN: 2303-0577 eISSN: 2830-7062 http://dx.doi.org/10.23960/jitet.v13i1.5876

IMPLEMENTASI ALGORITMA K-MEANS DALAM OPTIMALISASI PENGELOMPOKAN SURAT MASUK DI PG. SINDANGLAUT

Rosnita Amalia[1*], Nana Suarna[2], Irfan Ali[3], Dendy Indriya Efendi[4]

[1,2,3,4] STMIK IKMI Cirebon; Jl. Perjuangan No. 10B, Karyamulya, Kesambi, Kota Cirebon, Jawa Barat, Telp: 0231-490480

Received: 29 Desember 2024
Accepted: 14 Januari 2025
Published: 20 Januari 2025

Keywords:
Algoritma K-Means, Data Mining, Efisien Pengarsipan, Pengolahan Data.

Corespondent Email:
rosnitaamalia28@gmail.com

Abstrak. Pengelolaan surat masuk merupakan salah satu aspek penting dalam administrasi perusahaan, termasuk di PG. Sindanglaut yang berfungsi untuk mengelompokkan dan mengarsipkan surat secara efektif. Namun, dalam praktiknya, sistem pengelompokan surat yang digunakan saat ini masih manual dan kurang efisien, sehingga mengakibatkan keterlambatan dalam pengambilan keputusan dan pencarian dokumen. Untuk mengatasi permasalahan tersebut, penelitian ini melakukan penerapan Algoritma K-Means dalam mengoptimalkan proses pengelompokan surat masuk. Pemilihan Algoritma K-Means karena mampu melakukan *clustering* data secara cepat dan efisien, terutama dalam jumlah data yang besar. Permasalahan utama yang dihadapi adalah bagaimana meningkatkan efisiensi dan akurasi pengelompokan surat masuk menggunakan algoritma k-means untuk mendukung pengelolaan arsip yang lebih baik. Penelitian ini melalui metode eksperimen, di mana data surat masuk dianalisis dan dikelompokkan berdasarkan karakteristik kemiripannya menggunakan algoritma k-means sesuai dengan metode *Knowledge Discoveri in Database* (KDD). Dalam proses KDD mencakup data *selection*, *preprocessing*, *tranformation*, data mining untuk menentukan jumlah *cluster* optimal, dan *interpretation* atau evaluasi hasil *clustering* dengan mengukur nilai akurasi dan efektivitas pengelompokan berdasarkan nilai dari *Davies Bouldin Index* (DBI) dimana nilai terendah yang mendekati 0 menunjukkan klaster yang optimal. Tujuan penelitian ini adalah untuk mengoptimalkan sistem pengelompokan surat di PG. Sindanglaut, sehingga dapat meminimalkan waktu yang dibutuhkan untuk pencarian dokumen dan memperbaiki sistem pengarsipan. Penelitian ini menghasilkan bahwa penggunaan algoritma k-means dapat meningkatkan efisiensi pengelompokan surat masuk berdasarkan jenis surat yang terdiri dari Surat Dinas, Surat Permohonan, dan Surat Pemberitahuan sesuai dengan jenis dan kepentingannya, yang menghasilkan 8 klaster serta keakuratan pengelompokan yang sesuai dengan metode penelitian dimana nilai DBI yang diperoleh senilai 0.321.

Abstract. *The management of incoming mail is one of the important aspects of company administration, including at PG. Sindanglaut, which functions to effectively categorize and archive letters. However, in practice, the current mail categorization system is still manual and inefficient, resulting in delays in decision-making and document retrieval. To address these issues, this research implements the K-Means Algorithm to optimize the process of categorizing incoming mail. The choice of the K-Means Algorithm is due to its ability to perform data clustering quickly and efficiently, especially with large amounts of data. The main problem faced is how to improve the efficiency and accuracy of incoming mail categorization using the K-Means algorithm to*

support better archive management. This research employs an experimental method, where incoming mail data is analyzed and grouped based on similarity characteristics using the k-means algorithm in accordance with the Knowledge Discovery in Databases (KDD) method. The KDD process includes data selection, preprocessing, transformation, and data mining to determine the optimal number of clusters, as well as interpretation or evaluation of clustering results by measuring accuracy and effectiveness based on the Davies Bouldin Index (DBI), where the lowest value approaching 0 indicates optimal clusters. The aim of this research is to optimize the mail grouping system at PG. Sindanglaut, thereby minimizing the time required for document searches and improving the filing system. This study concludes that the use of the k-means algorithm can enhance the efficiency of grouping incoming mail based on types of letters, which consist of Official Letters, Application Letters, and Notification Letters according to their type and importance, resulting in 8 clusters and clustering accuracy consistent with the research method, where the obtained DBI value is 0.321.

1. PENDAHULUAN

Pengelompokan surat masuk yang tidak efisien dapat menyebabkan kesulitan dalam penelusuran dokumen, menghambat proses kerja, serta meningkatkan risiko kehilangan informasi penting. Dalam era digital saat ini, pengelolaan surat masuk di berbagai instansi dan perusahaan mengalami transformasi yang signifikan seiring dengan perkembangan teknologi informasi. [1] menyajikan sistem informasi yang dirancang untuk pencatatan surat menyurat secara efektif, menggunakan teknik *waterfall* dalam pengembangan sistem menunjukkan dampak positif dalam pengelolaan surat di instansi pemerintah, menawarkan solusi berbasis web yang mempermudah administrasi.

Berdasarkan hasil observasi yang dilakukan, permasalahan utama dalam pengelolaan surat masuk di berbagai instansi adalah efisiensi dan akurasi dalam proses pengelompokkan dan pengarsipan. Salah satu metode yang akan diimplementasikan adalah algoritma k-means teknik pengelompokan atau *clustering* yang dapat mengelompokkan data berdasarkan kemiripannya. [2] merancang sistem berbasis web untuk pengelolaan arsip yang lebih efisien, temuan penelitian ini dapat digunakan untuk mendukung pengembangan sistem dalam pengarsipan surat di masa depan, namun penelitian ini tidak menggunakan K-Means untuk mengelompokan surat.

Algoritma K-Means menjadi pilihan populer untuk menyelesaikan masalah *clustering* atau pengelompokan data, karena kemampuannya yang handal dalam menangani data dengan dimensi besar dan hasil yang *interpretable* [3]. K-Means salah satu teknik *partitional clustering* atau pengelompokan dengan sistem yang tidak memiliki tingkatan, yang digunakan untuk menganalisis data bersifat *machine learning* dan mengelompokkan data dengan sistem partisi [4]. [5] dalam penelitiannya mengaplikasikan k-means untuk klasterisasi kasus stunting balita, yang memberikan wawasan baru dalam pengelolaan data kesehatan masyarakat.

Dari beberapa peneliti terdahulu menjelaskan bahwa penelitiannya mengeksplorasi penerapan K-Means dalam mendukung pengelolaan koleksi di perpustakan, dengan analisis berbasis data peminjam dan respon pengguna [6]. Penelitian [7] dan [8] berfokus pada pengembangan sistem manajemen aplikasi berbasis web untuk mempermudah pengelolaan surat menyurat dan efisiensi dalam proses administasi, namun belum menggunakan teknik algoritma k-means untuk pengelompokkannya. [9] menerapkan algoritma k-means dalam mengelompokkan surat keluar di program studi, berfokus pada efisiensi pengelolaan surat di institusi pendidikan. Mereka menentukan bahwa k-means sangat baik dalam penggolongan data administratif, penelitian ini menguatkan bahwa penggunaan algoritma k-means dapat membantu mengorganisis dan mengelompokan surat masuk dengan kategori yang beragam.

Penelitian ini bertujuan untuk mengoptimalkan pengelompokan surat masuk di PG. Sindanglaut dengan menentukan teknik algoritma k-means, sehingga dapat

meningkatkan efisiensi dalam pengelolaan arsip dan mempermudah proses pencarian dokumen. Penelitian ini juga akan mengadopsi pendekatan eksperimen dan evaluasi dengan *Davies-Bouldin Index* (DBI) untuk mengukur efektivitas pengelompokan, dalam pengelompokan data menggunakan DBI untuk mengukur efisiensi *clustering* dimana nilai yang kurang dari 0 menunjukkan hasil *clustering* yang memiliki kualitas lebih baik dengan jarak antar *cluster* yang jelas [10].

Dengan implementasi penelitian ini, maka dapat memberikan kontribusi dalam peningkatan metode pengelompokan data arsip dengan menetapkan algoritma k-means. Implikasi penelitian ini juga dapat mengisi kesenjangan pengetahuan mengenai pengaplikasian algoritma k-means dalam konteks administrasi perkantoran dan bisa menjadi bahan referensi bagi peneliti lebih lanjut. Implementasi algoritma k-means ini dapat mengurangi beban kerja manual dan akurasi pengelompokkan surat masuk, sehingga dapat meningkatkan efisiensi operasional.

Metode yang diterapkan dalam penelitian ini adalah teknik kuantitatif sesuai prosedur *Knowledge Discovery in Databases* (KDD) dalam menerapkan k-means untuk mengelompokkan surat masuk. KDD adalah tata cara dalam mengenali dan mengubah dataset menjadi sebuah informasi [11].

Pendekatan ini dilakukan untuk menguji efektivitas dan efisiensi algoritma dalam mengelompokkan surat-surat yang masuk berdasarkan kategori yang sudah diatur.

2. TINJAUAN PUSTAKA

2.1 Surat Masuk

Surat merupakan suatu selembar kertas atau lebih yang tertulis suatu pernyataan atau informasi yang ingin disampaikan, diberitakan atau diutarakan kepada orang lain. Surat masuk ialah semua jenis surat yang diterima dari instansi lain maupun dari perorangan, termasuk yang diterima melalui kantor pos dengan mempergunakan buku pengiriman, dan surat dari kurir [12].

2.2 Clustering

Clustering merupakan proses pengelompokan data ke dalam kelompok-kelompok yang memiliki karakteristik atau atribut yang memiliki kemiripan [13]. Pengelompokan pada sekumpulan data yang terdiri dari anggota setiap data pada setiap klaster (k) dan unggul dari suatu data yang akan masuk ke dalam kelompok (*cluster*) berdasarkan kategori terpilih [9].

2.3 K-Means

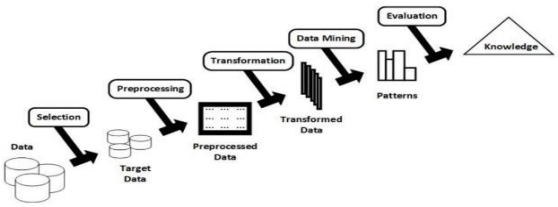
K-Means ialah algoritma yang digunakan dalam pengelompokkan data menjadi beberapa kelompok dengan beberapa klaster [14]. Algoritma ini bekerja dengan mengumpulkan data ke dalam kelompok (k) berdasarkan posisi titik pusat klaster [15]. Penerapannya adalah mengelompokan surat masuk berdasarkan beberapa fitur tertentu seperti jenis surat, tanggal surat, atau topik. Dengan data input ini, hasilnya akan diproses dan diuji berulang kali dengan beberapa iterasi untuk memperoleh hasil optimal. Oleh karena itu, surat-surat itu dikelompokan jadi beberapa kelompok secara otomatis.

2.4 DBI

Davies Bouldin Index (DBI) adalah ukuran yang akan menguji suatu pengelompokan yang merupakan metode yang diperkenalkan oleh David L. Davies dan Donald W. Bouldin. DBI ditentukan untuk mengevaluasi hasil *clustering* secara umum berdasarkan kuantitas dan kedekatan antar anggota *cluster*. Dalam menghitung nilai DBI dimana nilai yang semakin kecil maka *cluster* yang dihasilkan semakin baik [16].

3. METODE PENELITIAN

Teknik analisis data yang digunakan dalam penelitian ini adalah analisis *clustering* dengan Algoritma K-Means, yang merupakan metode non-hirarkis untuk mengumpulkan data beberapa jumlah *cluster* (k). Menganalisis data dengan pelaksanaan data mining menggunakan proses tahapan KDD. Tahapan ini diantaranya yaitu *selection data, preprocessing data, transformation data, data mining, dan evaluation data* [11].



Gambar 1. Tahapan KDD

3.1 Data Selection

Pada *data selection* ini peneliti mengambil data surat masuk berdasarkan jenis surat di PG. Sindanglaut dalam kurun tahun 2023 dengan format file excel.

3.2 Preprocessing Data

Langkah selanjutnya adalah melakukan *preprocessing* pada data, yaitu melakukan pembersihan data seperti mengatasi data yang hilang, menghapus atau memperbaiki data yang tidak sesuai, dan mengeliminasi beberapa atribut yang tidak diperlukan dalam perhitungan algoritma k-means *clustering*.

3.3 Data Transformation

Data yang sudah bersih akan diubah atau konversi ke format yang lebih sesuai dalam proses data mining. Transformasi ini meliputi proses normalisasi (mengubah data menjadi skala tertentu), tujuan dari transformasi data adalah untuk mebuktikan bahwa data siap digunakan dalam proses data mining.

3.4 Data Mining

Teknik ini digunakan untuk menggali informasi dari data. Pemilihan algoritma k-means digunakan untuk mengelompokkan data berdasarkan kemiripannya. Selain itu, untuk melakukan evaluasi kinerja menggunakan *Davies Bouldin Index* (DBI) untuk menetapkan jumlah (k) *cluster* yang ideal.

3.5 Evaluation

Pada tahap evaluasi, peneliti mengukur performa algoritma berdasarkan nilai evaluasi *Davies Bouldin Index* (DBI) dalam pengelompokan jenis surat masuk. Evaluasi nilai DBI yang paling rendah dilakukan untuk menentukan kualitas pengelompokan, yang mencakup kohesi antar anggota *cluster* dan keterpisahan antar *cluster*. Setelah itu, interpretasi dilakukan untuk mengidentifikasi pola yang muncul dari hasil *clustering*, yang akan digunakan sebagai dasar untuk memberikan rekomendasi dalam pengelolaan surat masuk.

3.6 Knowledge

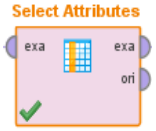
Tahap ini mengaitkan analisis menyeluruh terhadap tren, pola, atau data yang bertautan. Tujuan dari tahap ini adalah untuk membangun kesimpulan yang relevan dan membuat keputusan yang didukung oleh informasi yang diperoleh dari hasil telaah pada tahap sebelumnya.

4. HASIL DAN PEMBAHASAN

Berdasarkan hasil eksplorasi yang dijalankan menggunakan Algoritma K-Means untuk mengelompokkan jenis surat yang terbagi menjadi surat dinas, surat permohonan, dan surat pemberitahuan yang berdasarkan kesamaan karakteristik dari masing-masing jenis surat. Dalam proses ini peneliti menggunakan *software* Altair AI Studio versi 2024.1.0 adapun proses dalam pengelompokan yang berdasarkan metode *Knowledge Discovery in Database* (KDD) yang akan dijabarkan sebagai berikut:

4. 1 Data Selection

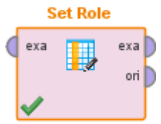
Dataset ini diperoleh dari instansi terkait yaitu PG. Sindanglaut dengan melakukan observasi, wawancara dan dokumentasi dengan staf administrasi dan mendapatkan data sebanyak 672 dataset yang memiliki sejumlah atribut, antara lain tanggal masuk surat, nomor agenda, asal surat, nomor surat, perihal, jenis surat, dan kelompok jumlah surat yang terbagi menjadi (surat dinas, surat permohonan, surat pemberitahuan).



Gambar 2. Operator Select Attribute

Selanjutnya menggunakan *operator Select Attribute* untuk menentukan atribut yang akan diolah dalam proses *clustering*. Pada dataset yang semula berjumlah 10 atribut, kemudian dipilih 4 atribut yang akan digunakan.

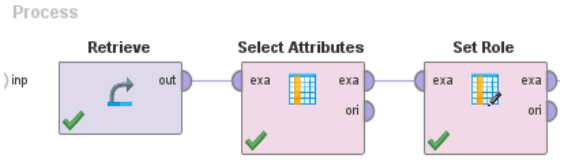
Kemudian memilih ID pada dataset menggunakan *operator Set Role* seperti pada gambar



Gambar 3. Operator Set Role

Pada tahap ini atribut yang akan dijadikan ID yaitu Jenis Surat dan Keperluan melalui *parameter set role* dimana *attributes name* memilih Jenis Surat dan Keperluan, dan target role memilih id.

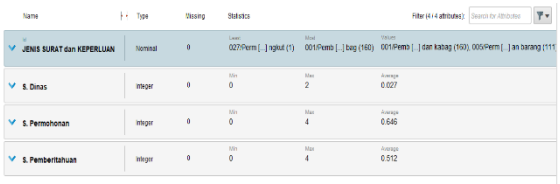
Proses pemodelan pada Altair AI Studio dalam data selection dapat di lihat pada gambar



Gambar 4. Proses Data Selection

4.2 Preprocessing

Pada penelitian ini tidak dilakukan *prepocessing* data untuk mengurangi jumlah data yang tidak relevan, karena pada dataset tidak ada data yang *missing* atau bernilai kosong dimana nilai dalam atribut S. Dinas, S. Permohonan, dan S. Pemberitahuan merupakan jumlah lembar surat berdasarkan jenis surat dalam satu nomor agenda. Adapun dataset yang sudah siap untuk di analisis dalam *K-Means Clustering* dapat dilihat pada gambar



Gambar 5. Data Statistik

Berdaraskan gambar yang ditampilkan, tidak didapati adanya *missing values* pada dataset. Dengan demikian, proses analisis dapat dilanjutkan ke tahap selanjutnya.

4.3 Transformation

Proses *transformation* data ke bentuk yang sesuai untuk diproses dalam data mining. Tujuan utama dari transformasi data agar lebih efektif dan efisien dalam pengolahan dan analisis yang diproses oleh algoritma yang digunakan. Pada penelitian ini tahap *transformation* tidak diperlukan karena data telah berbentuk numerik dan siap diolah pada Altair AI Studio menggunakan algoritma k-means.

4. 4 Data Mining

Melakukan data mining dengan memastikan jumlah *cluster* merupakan hal yang penting untuk mendapatkan hasil yang akurat. Nilai klaster atau (k) yang telah ditentukan berdasarkan nilai *cluster* yang optimal dan jumlah anggota yang optimal dalam tiap *cluster*. Proses ini menggunakan operator K-Means Clustering untuk pemodelan klasterisasi seperti pada gambar

Gambar 6. Operator K-Means Clustering

Peneliti menggunakan nilai K=8 pada parameter operator K-Means Clustering dengan jumlah *Max Runs* 10, dan *Measure Types Numerical Measures* dengan jenis *Euclidean Distance*. Dalam algoritma k-means pemilihan *Measure Types* atau metode pengukuran sangat penting untuk menentukan cara algoritma menghitung jarak atau kesamaan antar data untuk mengelompokkan ke dalam *cluster*. Peneliti memilih *Numerical Measures* karena data sudah dikonversi menjadi format numerik, dalam *Numerical Measure* memilih *Euclidean Distance* yang dimana *Euclidean Distance* digunakan untuk menghitung jarak antara titik pada data dengan mempertimbangkan perbedaan absolut dalam nilai-nilai numerik untuk mengelompokkan titik-titik yang memiliki kesamaan nilai. Pada tahap *clustering* ini diperoleh nilai keanggotaan setiap *cluster* seperti yang ditampilkan dalam tabel.

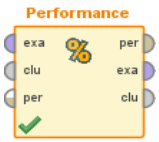
Tabel 1. Nilai Anggota Cluster

Cluster	**Jumlah Anggota**
Cluster_0	21 *items*
Cluster_1	375 *items*
Cluster_2	213 *items*
Cluster_3	2 *items*

Cluster_4	2 *items*
Cluster_5	24 *items*
Cluster_6	28 *items*
Cluster_7	7 *items*
Total *number of items*	672 *items*

4.5 Interpretation/Evaluation

Selanjutnya lakukan evaluasi dari hasil *K-Means Clustering* menggunakan operator *Cluster Distance Performance* dipakai dalam mengukur performa dari operator *K-Means Clustering* dengan menggunakan teknik *Davies Bouldin Index* (DBI) tujuannya untuk mengetahui nilai DBI dari hasil proses *clustering*.



Gambar 7. Operator Clustering Distance Performance

Setelah operator *Cluster Distance Performance* dijalankan dengan parameter *Main Criterion* pada *Davies Bouldin* mendapatkan informasi nilai DBI sebesar 0.321 menunjukkan *cluster* yang optimal berada pada K=8 yang terdiri dari C0, C1, C2, C3, C4, C5, C6, dan C7 sehingga K=8 digunakan dalam penelitian ini.

Mengevaluasi kualitas hasil *clustering* yang dilakukan berdasarkan akurasi nilai DBI dengan metode *Algoritma K-Means Clustering* menggunakan *Measure Types Numerical Measures* dengan jenis *Euclidean Distance*, ditemukan bahwa *cluster* yang optimal dengan nilai yang mendekati 0 berada pada K=8 dengan nilai DBI 0.321. Dimana nilai DBI yang lebih kecil memperlihatkan klaster yang lebih baik, karena jarak antar-klaster lebih besar dan jarak dalam klaster lebih kecil. Secara keseluruhan, hasil dari pengelompokan data surat masuk menggunakan Algoritma K-Means yang dievaluasi berdasarkan nilai *Davies Bouldin Index* (DBI) memberikan gambaran kualitas keefektivan *clustering* data berdasarkan kedekatan dengan centroid pada masing-masing klaster.

4.6 Knowledge

Hasil analisis menunjukkan bahwa klasterisasi berhasil menghasilkan *cluster* dengan kualitas terbaik pada K=8, atau 0.321, dengan nilai DBI paling mendekati 0.

Gambar 8. Hasil Clustering dalam Scater Bubble

Gambar tersebut menunjukkan penyebaran dari tiga jenis surat yaitu "S. Pemberitahuan," "S. Dinas," dan "S. Permohonan," berdasarkan hasil proses clustering. Pada sumbu vertikal menunjukkan jumlah dari kategori jenis surat dengan nilai berkisaran dari 0 hingga 5, sementara pada sumbu horizontal terdapat jenis surat yang dianalisis dengan kode unik seperti 001, 002, 003, hingga 027 dengan deskripsi keperluan surat yang menunjukkan kategori kelompok hasil clustering. Setiap jenis surat di kelompokkan berdasarkan keperluan tertentu, seperti permohonan perbaikan, pemberitahuan untuk kabag, dan kebutuhan administratif lainnya.

Penyebaran Surat Pemberitahuan ditandai dengan warna oranye, dimana jenis surat ini menunjukkan tingginya penggunaan surat pemberitahuan dibandingkan kategori lainnya. Penyebaran Surat Dinas ditandai oleh warna biru yang memiliki nilai paling sedikit di beberapa titik, titik data pada kategori ini sebagian besar berada pada nilai rendah di sumbu vertikal. Kemudian pada penyebaran S. Permohonan yang berwarna hijau yang terlihat penyebarannya merata pada berbagai jenis surat yang ditunjukkan dari variasi dalam keperluannya.

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa setiap jenis surat memiliki pola penyebaran yang berbeda pada setiap cluster. Dari hasil *clustering* ini di dominasi oleh Surat Pemberitahuan yang memiliki peran penting dalam administratif dengan sifat surat yang cenderung digunakan untuk

menyampaikan informasi atau keputusan penting. Sebaliknya, Surat Dinas memiliki tingkatan yang lebih rendah mencerminkan penggunaan surat lebih spesifik pada situasi tertentu seperti pengambilan keputusan dalam organisasi. Persebaran yang merata ditunjukkan oleh Surat Permohonan yang digunakan untuk berbagai keperluan yang mencakup permintaan sumber daya, izin, atau klarifikasi tertentu.

5. KESIMPULAN

a. Berdasarkan hasil penelitian pada analisis data surat masuk di PG. Sindanglaut dan pembahasan yang dilakukan, dapat disimpukan bahwa penelitian ini berhasil mengidentifikasi pola intensitas penggunaan surat pada beberapa kategori, yaitu S. Permohonan, S. Pemberitahuan, dan S. Dinas. Melalui proses *clustering* dengan Algoritma K-Means, penelitian ini mampu mengelompokkan jenis-jenis surat berdasarkan pola dan frekuensinya. Hasil *clustering* menunjukkan perbedaan yang signifikan dalam intensitas penggunaan surat pada setiap kategori dan memberikan informasi penting terkait pengelolaan administrasi.
b. Penelitian ini menghasilkan nilai *Davies-Bouldin Index* (DBI) sebesar 0.321 menunjukkan bahwa hasil *clustering* memiliki kualitas pemisahan *cluster* yang baik. Semakin rendah nilai DBI semakin baik kualitas *cluster* yang terbentuk, yang berarti setiap *cluster* memiliki kemiripan karakteristik dalam kelompoknya masing-masing. Hal ini relevan dengan tujuan penelitian untuk mendapatkan *cluster* yang optimal dan memberikan gambaran jelas tentang pola intensitas penggunaan surat di setiap kategori.
c. Penelitian ini berhasil mencapai tujuan utama dalam menganalisis intensitas penggunaan surat masuk dengan menerapkan metode *K-Means Clustering*. Temuan yang dihasilkan dapat digunakan sebagai dasar untuk pengelolaan administrasi surat yang efisien dan efektif. Selain itu, struktur yang diterapkan dalam penelitian ini terbukti mampu memberikan hasil *clustering* yang optimal, sehingga dapat digunakan untuk penelitian lebih lanjut yang melibatkan data dengan karakteristik yang serupa.

UCAPAN TERIMA KASIH

Dalam proses penelitian ini penulis mengucapkan terima kasih bagi semua pihak yang telah berpartisipasi.

DAFTAR PUSTAKA

[1] A. B. Praja, D. Darmansah, and S. Wijayanto, "Sistem Informasi Pencatatan Surat Masuk dan Surat Keluar Berbasis Website Menggunakan Metode Waterfall," *J. Sist. Komput. dan Inform.*, vol. 3, no. 3, p. 273, 2022, doi: 10.30865/json.v3i3.3914.

[2] D. H. L. Dara and Liza fitria, "Sistem Informasi Pengelolaan Arsip Surat Berbasis Web di Kantor Badan Pertanahan Nasional Kota Langsa," *J-ICOM - J. Inform. dan Teknol. Komput.*, vol. 2, no. 2, pp. 62–69, 2021, doi: 10.33059/j-icom.v2i2.3795.

[3] F. Nuraeni, H. Susilawati, and Y. Handoko Agustin, "Perbandingan Implementasi Algoritma K-Means++ Dan Fuzzy C-Means Pada Segmentasi Citra Wajah," *JuTI "Jurnal Teknol. Informasi,"* vol. 1, no. 2, p. 47, 2023, doi: 10.26798/juti.v1i2.722.

[4] Abdussalam Amrullah, Intam Purnamasari, Betha Nurina Sari, Garno, and Apriade Voutama, "Analisis Cluster Faktor Penunjang Pendidikan Menggunakan Algoritma K-Means (Studi Kasus: Kabupaten Karawang)," *J. Inform. dan Rekayasa Elektron.*, vol. 5, no. 2, pp. 244–252, 2022, doi: 10.36595/jire.v5i2.701.

[5] P. Apriyani, A. R. Dikananda, and I. Ali, "Penerapan Algoritma K-Means dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi," *Hello World J. Ilmu Komput.*, vol. 2, no. 1, pp. 20–33, 2023, doi: 10.56211/helloworld.v2i1.230.

[6] I. Padiku and A. Lahinta, "Penerapan Clustering K-Means Untuk Mendukung Pengelolaan Koleksi Pada Perpustakaan Fakultas Teknik Universitas Negeri Gorontalo," *J. Tek.*, vol. 20, no. 1, pp. 54–62, 2022, doi: 10.37031/jt.v20i1.206.

[7] D. Nurdiana, "Implementasi Aplikasi Pengelolaan Surat Masuk Dan Keluar Berbasis Web Di Prodi Sistem Informasi," *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 6, no. 2, pp. 135–144, 2020, doi: 10.33330/jurteksi.v6i2.437.

[8] Riswandi Ishak, Setiaji, Fajar Akbar, and Mahmud Safudin, "Rancang Bangun Sistem Informasi Surat Masuk Dan Surat Keluar Berbasis WEB Menggunakan Metode Waterfall," *J. Indones. Sos. Teknol.*, vol. 1, no. 3, pp. 198–209, 2020, doi:

10.36418/jist.v1i3.33.
[9] L. Rusdiana and V. C. Hardita, "Algoritma K-Means dalam Pengelompokan Surat Keluar pada Program Studi Teknik Informatika STMIK Palangkaraya," *J. SAINTEKOM*, vol. 13, no. 1, pp. 55–66, 2023, doi: 10.33020/saintekom.v13i1.356.
[10] M. Wahyudi, S. Solikhun, and L. Pujiastuti, "Komparasi K-Means Clustering dan K-Medoids Clustering dalam Mengelompokkan Produksi Susu Segar di Indonesia Berdasarkan Nilai DBI," *J. Bumigora Inf. Technol.*, vol. 4, no. 2, pp. 243–254, 2022, doi: 10.30812/bite.v4i2.2104.
[11] F. Zahra, A. Khalif, B. N. Sari, U. S. Karawang, T. Timur, and J. Barat, "PROVINSI INDONESIA MENGGUNAKAN ALGORITMA K-MEDOIDS," vol. 12, no. 2, 2024.
[12] R. R. S. Anawoli, A. A. Pekuwali, and P. A. R. L. Lede, "System Development dalam Pengarsipan Surat Berbasis Model Object Oriented Programming," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 2, pp. 463–471, 2024, doi: 10.57152/malcom.v4i2.1232.
[13] B. Putra Aryadi and N. Hendrastuty, "Penerapan Algoritma K-Means Untuk Melakukan Klasterisasi Pada Varietas Padi," *J. Inform. Rekayasa Elektron.*, vol. 7, no. 1, pp. 124–129, 2024, [Online]. Available: http://e-journal.stmiklombok.ac.id/index.php/jireISSN .2620-6900
[14] N. Dwitri, J. A. Tampubolon, S. Prayoga, and P. P. P. A. N. W. F. I. R. H. Zer, "PENERAPAN ALGORITMA K-MEANS DALAM MENENTUKAN TINGKAT PENYEBARAN PANDEMI COVID-19 DI INDONESIA," vol. 4, no. 1, pp. 128–132, 2020.
[15] Y. Hasan, U. B. Darma, K. Medan, and S. Utara, "PENGUKURAN SILHOUETTE SCORE DAN DAVIES- BOULDIN INDEX PADA HASIL CLUSTER K-MEANS DAN," vol. 12, no. 3, 2024.
[16] T. Hardiani, "Analisis Clustering Kasus Covid 19 di Indonesia Menggunakan Algoritma K-Means," *J. Nas. Pendidik. Tek. Inform.*, vol. 11, no. 2, pp. 156–165, 2022, doi: 10.23887/janapati.v11i2.45376.