

**“APLIKASI SURAT MASUK DAN KELUAR DENGAN KLASTERISASI
MENGUNAKAN METODE *K-MEANS* DI INSTALASI FARMASI RSUD
DR.SAIFUL ANWAR”**

Sari Nur Sita Wibowo¹

Amak Yunus EP²

¹Sistem Informasi, Universitas Kanjuruhan Malang, sitha.wibowo@gmail.com

²Sistem Informasi, Universitas Kanjuruhan Malang, amakyunus@yahoo.com

ABSTRAK

Surat menyurat adalah suatu kegiatan untuk mengadakan hubungan secara terus menerus antara pihak yang satu kepada pihak yang lainnya, dan dilaksanakan dengan saling berkirim surat. Di era teknologi yang semakin berkembang, pemanfaatan teknologi sudah menjadi aspek penting dalam kehidupan manusia. Salah satu bidang yang terkena dampaknya yaitu dalam kegiatan surat masuk dan keluar dalam suatu organisasi.

Oleh karena itu perlu adanya membangun suatu aplikasi surat masuk dan keluar, yang mana dapat memudahkan pekerjaan manusia dalam mengirimkan surat, mengarsipkan surat, dan mengelompokkan surat sesuai dengan bagian yang ada didalam organisasi. Aplikasi disajikan dengan klasterisasi menggunakan metode *K-Means*, yang mana metode ini dapat digunakan untuk dataset yang besar maupun dataset kecil.

Kata Kunci : Surat Masuk dan Keluar, Klasterisasi, *K-Means*.

ABSTRACT

Correspondence is an activity to arrange relationship continuously between one to another, and be held with sending a mail each other. In this technolgy era which develop increasingly, technology utilization has become important aspect in human life. One of sector which affected is in activity of incoming or outcoming mail on an organization.

*Therefore, it is necessary to build an application incoming and outgoing mail, which can facilitate the work of human in sending mail, archiving mail and classifying the mail corresponding related to the organization. Applications presented by clustering using *K-Means*, therefore this method can be used for large datasets or a small dataset.*

*Keywords: Inbox and Outbox, Clustering, *K-Means**

1. Pendahuluan

Dalam pengelolaan surat masuk dan surat keluar pada suatu instansi termasuk kegiatan dalam mengolah data di dalam komputer yang dalam pengelolaan surat masuk dan surat keluar tersebut harus dilakukan oleh suatu instansi atau organisasi. Kegiatan surat menyurat harus mendapatkan perhatian yang lebih, karena isi dari surat instansi akan menjadi sarana pencapaian instansi. Selama ini proses surat masuk dan surat keluar di instalasi masih di kirim dengan cara manual, yaitu staf mengirimkan surat ke bagian lain dan diarsipkan dengan cara dicatat kedalam buku besar. Proses yang dilakukan selama ini tidak dipisahkan berdasarkan bagian yang ada didalam instansi. Hal ini menimbulkan kesulitan di dalam pencarian data dari surat masuk dan surat keluar jika suatu waktu diperlukan, serta menimbulkan keterlambatan tersampainya surat secara langsung. Oleh karena itu, diperlukan adanya sistem yang dapat membantu dalam proses pengiriman, pencarian, serta pengarsipan surat, agar dapat mempercepat dalam proses pengiriman dan pencarian surat di dalam suatu instansi.

2. Tinjauan Pustaka

2.1 Konsep Pengelompokkan

Dalam pengelompokkan, data-data yang datang tidak mempunyai label kelas, tetapi kemudian dikelompokkelompokkan menurut karakteristiknya, barulah label tersebut dapat diberi label sesuai hasil karakteristik kelompok masing-masing. (Emha, 2009)

2.2 Tujuan Pengelompokkan

Tujuan pekerjaan pengelompokkan (*clustering*) data dapat dibedakan menjadi dua, yaitu pengelompokkan untuk pemahaman dan pengelompokkan untuk penggunaan. Jika tujuannya untuk pemahaman, kelompok yang terbentuk harus menangkap struktur alami data, biasanya proses pengelompokkan dalam tujuan ini hanya sebagai proses awal

untuk kemudian dilanjutkan dengan pekerjaan inti seperti peringkasan atau *summarization* (rata-rata, standard deviasi), pelabelan kelas pada setiap kelompok untuk kemudian digunakan sebagai data latih klasifikasi dan sebagainya. Sementara jika untuk penggunaan, tujuan utama pengelompokkan biasanya adalah mencari *prototype* kelompok yang paling representatif terhadap data, memberikan abstraksi dari setiap objek data dalam kelompok dimana sebuah data terletak di dalamnya.

2.3 K-Means

K-Means merupakan salah satu metode pengelompokkan data nonhierarki (sekatan) yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan ke dalam kelompok yang lain (Yudi Agusta, 2007). Adapun langkah-langkah dalam metode k-means adalah sebagai berikut :

- Pilih jumlah *cluster* k .
- Inisialisasi k pusat *cluster* ini bisa dilakukan dengan cara random.
- Hitung jarak semua data kesetiap titik pusat *cluster* dapat menggunakan teori jarak *Euclidean* yaitu sebagai berikut :

$$D(i, j) = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2}$$

Dimana :

$D(i, j)$ = Jarak data ke i ke pusat *cluster* j

X_{ik} = Data ke i pada atribut data ke k

X_{jk} = Titik pusat ke j pada atribut ke k .

- Menentukan titik pusat *cluster* baru.

- e. Hitung kembali setiap data dengan titik pusat *cluster* baru. Jika pusat *cluster* tidak berubah lagi maka proses *clustering* selesai.

3. Pembahasan

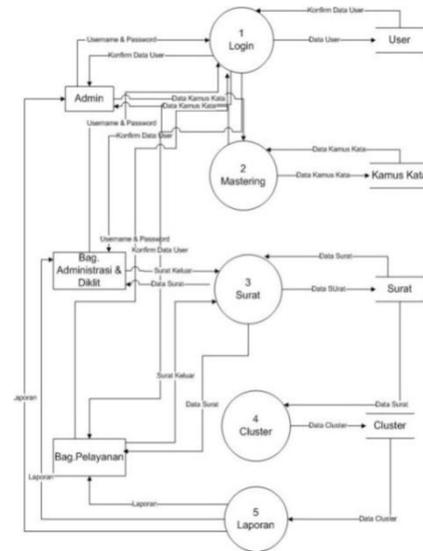
3.1 Desain Sistem

Diagram konteks terdapat bagian-bagian penting yaitu beberapa entitas, proses, dan arus data. Berikut adalah diagram konteks pada sistem yang akan dibangun.



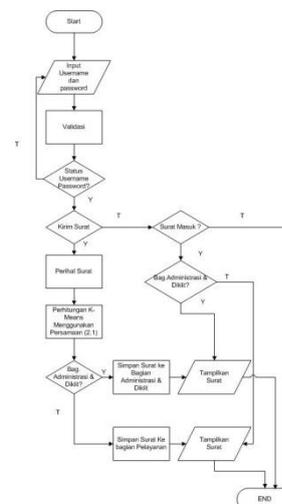
Gambar 1 Diagram Konteks

Secara garis besar dalam DFD Level 1 mempunyai lima proses yang merupakan proses utama pada sistem, yaitu admin dan *user* dapat mengakses *login*, kemudian admin dapat melakukan *setup* master, kemudian *user* dapat menginputkan surat keluar, yang kemudian surat diproses di dalam pengklusteran dengan menggunakan perihal surat yang ada. Setelah proses pengklusteran selesai, maka surat dapat dikirimkan ke bagian yang diinginkan, kemudian akan menjadi suatu laporan yang mana laporan tersebut dikelompokkan berdasarkan dengan bagian yang ada. Dan laporan tersebut dapat dilihat baik *user* maupun admin. Adapun DFD Level 1 disajikan pada gambar 2.



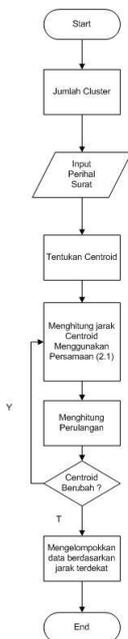
Gambar 2 DFD Level 1

flowchart sistem dimulai dari start setelah itu proses *login*, jika *login* sukses maka dapat menginputkan data. Setelah itu data akan tersimpan di data surat berdasarkan bidang yang ada. Jika proses *login* gagal, maka proses kembali ke awal. *Flowchart* sistem disajikan pada gambar 3.



Gambar 3 Flowchart Sistem

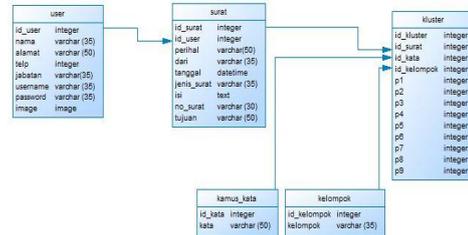
Proses *flowchart* klasterisasi K-Means dimulai dari start setelah itu proses menentukan jumlah cluster, kemudian menginputkan perihal surat melalui surat keluar yang akan dikirim. Setelah itu tentukan centroid dan menghitung jarak centroid, kemudian hitung perulangan, jika nilai centroid berubah maka menghitung kembali jarak centroid, jika nilai centroid tidak berubah maka dapat dikelompokkan berdasarkan jarak terdekat. Proses algoritma disajikan dalam Gambar 4.



Gambar 4 Flowchart K-Means Clustering

Gambar 5 merupakan Physical Data Model (PDM), Yang mana relasi tabel user dengan tabel surat menunjukkan relasi *one to many*, relasi tabel surat dengan tabel *cluster* menunjukkan relasi *one to one*, relasi tabel kamus kata dengan tabel *cluster* menunjukkan relasi *one to many*, relasi tabel kelompok dengan tabel

cluster menunjukkan relasi *one to many*.



Gambar 5 Physical Data Model

3.2 Pengujian Perhitungan Manual K-Means

a. Pembobotan Kata

Keyword perihal surat harus dibobotkan terlebih dahulu, karena k-means memproses hanya dapat memproses angka. Pembobotan kata disajikan dalam gambar 5.

Tabel 1 Pembobotan Kata

No	A1	A2	A3	A4	A5	A6	A7	A8	A9
1	60	80	70	60	80	50	80	70	70
2	70	60	80	60	70	60	80	60	60
3	70	70	60	80	80	80	60	70	80
4	60	80	70	60	80	50	80	70	70
5	70	60	80	60	70	60	80	60	60
6	70	70	60	80	80	80	60	70	80
7	50	70	70	60	70	70	60	70	60
8	60	50	70	60	70	70	70	70	60
9	50	50	70	60	70	60	70	80	80
10	60	70	80	70	60	80	70	60	70
11	50	50	70	70	80	60	60	80	60
12	60	60	60	70	60	70	70	60	60
13	70	60	70	50	80	80	80	80	80
14	80	70	70	80	80	70	70	70	60
15	60	70	70	70	80	80	70	60	50
16	30	50	50	40	50	50	30	40	30
17	70	60	70	50	80	80	80	80	80
18	60	70	80	70	60	80	70	60	70
19	30	40	50	40	40	40	30	40	30
20	30	30	30	30	30	30	30	80	40

b. Menentukan Jumlah Cluster

Data perihal surat merupakan aspek penting untuk mengetahui pengelompokan data. Pada perhitungan berikut ini mengelompokkan sebanyak 2 *Centroid* terkait penempatan surat berdasarkan ruangan. Ruangan 1 atau *Centroid* 1 merupakan bagian Administrasi & Diklit sedangkan

Centroid 2 merupakan bagian Pelayanan.

c. Menentukan Nilai Centroid

Untuk menentukan *centroid* awal (*initial centroid*), metode yang digunakan adalah mengambil data dari data sumber, secara acak atau random. Adapun nilai *centroid* disajikan dalam tabel 2.

Tabel 2 Nilai Centroid

Centroid	A1	A2	A3	A4	A5	A6	A7	A8	A9
C1	60	80	70	60	80	50	80	70	70
C2	30	40	50	40	40	40	30	40	30

d. Menghitung Jarak Data Kesetiap Titik Pusat Centroid

Menghitung jarak data ke setiap titik pusat *centroid*, dengan menggunakan rumus *euclidean*. Setelah melakukan perhitungan jarak data dengan rumus *Euclidean*, kemudian mengelompokkan data sesuai dengan jarak terpendek. Jika $d(x_1, c_1) < d(x_2, c_1)$ maka X_1 masuk pada *cluster* Administrasi dan jika $(x_2, c_1) > d(x_1, c_1)$ maka X_2 masuk pada *cluster* Pelayanan. Adapun hasil perhitungan jarak data ke setiap titik pusat centroid disajikan dalam tabel 3.

Tabel 3 Hasil Jarak Centroid

Surat	DC1	DC2	Administrasi	Pelayanan
1	0	100	ok	
2	31.62278	91.6515139	ok	
3	45.82576	108.1665383	ok	
4	0	100	ok	
5	31.62278	91.6515139	ok	
6	45.82576	108.1665383	ok	
7	34.64102	81.24038405	ok	
8	40	83.66600265	ok	
9	38.72983	91.10433579	ok	
10	42.42641	95.91663047	ok	

Tabel 3 (Lanjutan)

Surat	DC1	DC2	Administrasi	Pelayanan
11	42.42641	84.85281374	ok	
12	41.23106	80.62257748	ok	
13	41.23106	110.9053651	ok	
14	38.72983	106.3014581	ok	
15	41.23106	93.27379053	ok	
16	92.19544	17.32050808		ok
17	41.23106	110.9053651	ok	
18	42.42641	95.91663047	ok	
19	100	0		ok
20	110.9054	50		ok

e. Pengulangan (Menentukan Nilai Centroid)

Menghitung atau menentukan *centroid* baru. *Centroid* baru dihitung atau ditentukan dengan menghitung nilai rata-rata data pada setiap *cluster*. Adapun nilai *centroid* baru disajikan dalam tabel 4.

Tabel 4 Centroid Baru

Centroid	A1	A2	A3	A4	A5	A6	A7	A8	A9
C1	60	80	70	60	80	50	80	70	70
C2	30	40	50	40	30	50	40	40	30

f. Hitung Jarak Data Dengan Centroid Baru

Menghitung jarak setiap data kembali dengan centroid baru yang menggunakan rumus *euclidean*. Setelah melakukan perhitungan, kemudian langsung mengelompokkan data sesuai dengan jarak terpendek. Jika $d(x_1, c_1) < d(x_2, c_1)$ maka X_1 masuk pada *cluster* Administrasi dan jika $(x_2, c_1) > d(x_1, c_1)$ maka X_2 masuk pada *cluster* Pelayanan.

Tabel 5 Data Cluster Pengulangan

Surat	DC1	DC2	Administrasi	Pelayanan
1	0	99.49874371	ok	
2	31.62278	88.88194417	ok	
3	45.82576	106.7707825	ok	
4	0	99.49874371	ok	
5	31.62278	88.88194417	ok	

Tabel 5 (Lanjutan)

Surat	DC1	DC2	Administrasi	Pelayanan
6	45.82576	106.7707825	ok	
7	34.64102	77.45966692	ok	
8	40	80.62257748	ok	
9	38.72983	89.4427191	ok	
10	42.42641	91.10433579	ok	
11	42.42641	85.44003745	ok	
12	41.23106	76.15773106	ok	
13	41.23106	107.7032961	ok	
14	38.72983	104.8808848	ok	
15	41.23106	90.55385138	ok	
16	92.19544	24.49489743		ok
17	41.23106	107.7032961	ok	
18	42.42641	91.10433579	ok	
19	100	52.91502622		ok
20	110.9054	52.91502622		ok

Karena *centroid* tidak mengalami perubahan (sama dengan *centroid* sebelumnya) maka proses *Clustering* selesai.

4. Kesimpulan

Dari penelitian yang telah dilakukan, maka diperoleh kesimpulan : Berdasarkan hasil pengujian, sistem yang ada dengan sistem yang baru memiliki perbedaan kecepatan waktu dalam proses pengiriman dan pencarian surat. Yang mana dengan menggunakan sistem yang baru lebih efisien atau lebih cepat untuk melakukan pengiriman dan pencarian surat dibandingkan dengan sistem yang ada

5. Saran

Saran untuk penelitian pengembangan selanjutnya adalah sebagai berikut :

1. Aplikasi dapat dikembangkan dengan jangkauan yang lebih . luas,

dengan digunakan di seluruh Instalasi RSUD Dr.Saiful Anwar.

2. Aplikasi dapat dikembangkan dengan pengembangan berbasis *Mobile* agar lebih mudah digunakan.

6. Daftar Pustaka

Agusta Y, 2007, *K-means* - Penerapan, Permasalahan dan Metode Terkait. *Jurnal Sistem dan Informatika*, Vol. 3, 2007, Bali.

Ong J Oscar, 2013, Implementasi Algoritma K-Means Clustering untuk Menentukan Strategi Marketing President University. *Jurnal Ilmiah Teknik Industri*, Vol. 12, No. 1, ISSN 1412-6869, 2013, Bekasi.

Emha Taufiq Lutfi dan Kusri, 2009, *Algoritma Data Mining*. Andi, 2009, Yogyakarta.