

Sistem Segmentasi Keluhan Pelanggan di Perumda Air Minum Tirta Raharja Cimahi Menggunakan Metode K-Medoids

Febry Ramadhan¹, Yulison Herry Chrisnanto², Ade Kania Ningsih³

^{1,2,3}Jurusan Informatika, Fakultas Sains dan Informatika, Universitas Jenderal Achmad Yani, Cimahi, Indonesia

e-mail: febryramadhan285@gmail.com^{*1}, y.chrisnanto@gmail.com², kania215@gmail.com³

INFORMASI ARTIKEL

Sejarah Artikel:

Diterima Redaksi : 23 Maret 2021

Revisi Akhir : 25 November 2021

Diterbitkan Online : 25 November 2021

Kata Kunci:

K-Medoids, Keluhan, Clustering, Segmentasi, Perumda Tirta Raharja Cimahi

Korespondensi:

Telepon / Hp : +62 896-4816-5065

E-mail : febryramadhan285@gmail.com

A B S T R A K

Perumda Air Minum Tirta Raharja Cimahi merupakan sebuah perusahaan yang bergerak di bidang jasa pelayanan pengelolaan air minum dan pengelolaan sarana air limbah di Cimahi dan sebagian Kabupaten Bandung Barat, untuk meningkatkan kesejahteraan masyarakat yang mencakup aspek ekonomi, sosial, kesehatan dan pelayanan umum, namun dalam prosesnya tentu saja perusahaan banyak menerima keluhan dari konsumen yang berasal dari berbagai golongan pelanggan dengan jenis keluhan yang berbeda – beda, dari keluhan – keluhan tersebut tentu ada kecenderungan jenis keluhan yang ada pada suatu wilayah tertentu. Masalah yang ada meliputi tidak mengalirnya air, air yang berbau, dan lain sebagainya. Pendekatan yang digunakan untuk menyelesaikan masalah pada penelitian ini adalah menggunakan K-Medoids clustering. K-Medoids adalah teknik partisi yang berfungsi untuk mengelompokkan data set dari n objek kedalam kelompok k yang dikenal apriori. Dibandingkan dengan K-Means, K-Medoids lebih kuat untuk mengatasi kebisingan (noise) dan pencilan (outlier) karena meminimalkan sejumlah dissimilarities berpasangan, bukan jumlah kuadrat jarak Euclidean.

1. PENDAHULUAN

Keluhan merupakan suatu umpan balik dari produk atau jasa yang digunakan oleh pelanggan kepada produsen atau penyedia jasa yang bersangkutan. Semakin banyak keluhan yang datang artinya dibutuhkan adanya perhatian ekstra bagi setiap perusahaan yang menerima keluhan untuk memperbaiki atau meningkatkan kualitas produk atau jasa [1]. Keluhan itu sendiri pasti pernah diterima oleh setiap perusahaan, namun terkadang perusahaan kurang mengerti mengenai masalah yang dikeluhkan oleh konsumen padahal keluhan konsumen adalah hal yang sangat penting, karena jika tidak ditangani dengan baik akan berpengaruh pada loyalitas konsumen dan citra perusahaan itu sendiri. Perilaku keluhan sendiri sangat erat kaitannya dengan kepuasan konsumen yang menjadi tujuan utama semua perusahaan karena akan mempengaruhi pembelian ulang, kesetiaan, dan pemberitaan baik tentang perusahaan.

Berdasarkan hal tersebut pihak Perumda Air Minum Tirta Raharja mengalami kesulitan untuk memetakan kecenderungan masalah yang terjadi, sehingga berdampak pada strategi peningkatan mutu pelayanan dari Perumda Air Minum Tirta Raharja itu sendiri. Sebelumnya perusahaan belum melakukan teknik pengelompokkan secara khusus untuk memetakan kelompok jenis keluhan dan wilayah keluhan berasal, sehingga seringkali apa yang perusahaan lakukan untuk menangani masalah tersebut justru tidak sampai ke akar permasalahannya sehingga keluhan yang serupa kembali muncul dan akhirnya perusahaan mengeluarkan waktu, tenaga, dan biaya yang lebih untuk menyelesaikan masalah yang relatif serupa. *Data mining* merupakan proses analitik yang dirancang untuk meninjau sejumlah

data yang besar dalam mencari suatu pengetahuan tersembunyi yang berharga dan konsisten [2]. *Data mining* juga ialah proses pencarian pola yang menarik dan tersembunyi dari suatu kumpulan data yang berdimensi besar yang tersimpan dalam suatu basis data. [3]. Pada penelitian ini *data mining* akan digunakan untuk mengelola data pengaduan atau keluhan pelanggan di Perumda Air Minum Tirta Raharja.

Perumda Air Minum Tirta Raharja merupakan perusahaan yang bergerak di bidang pengelolaan air minum dan pengelolaan sarana air limbah di daerah, untuk meningkatkan kesejahteraan masyarakat yang mencakup berbagai aspek seperti ekonomi dan lainnya.

Terdapat beberapa teknik untuk melakukan klusterisasi, diantaranya K-Means, K-Medoids dan lainnya. Berdasarkan penelitian yang dilakukan oleh (Aishwaya) didapati kesimpulan bahwa algoritma K-Means lebih efisien untuk kumpulan data yang berjumlah kecil sementara K-Medoids nampaknya mempunyai performa yang lebih baik untuk dataset besar [4].

Berdasarkan pada uraian yang telah dijelaskan sebelumnya, maka dapat dirumuskan permasalahan yang menjadi fokus penelitian yaitu, banyaknya jumlah keluhan yang masuk menyebabkan proses analisis data menjadi sulit karena harus mengkombinasikan jenis keluhan dengan wilayah tersebar yang masuk dalam wilayah pelayanan perusahaan, sehingga diperlukan sebuah perangkat lunak yang tepat untuk membantu proses analisis data menjadi lebih cepat dan efisien.

Berdasarkan pada permasalahan yang telah diuraikan, tujuan dari penelitian ini adalah untuk mengetahui kelompok - kelompok data yang merepresentasikan keluhan berdasarkan wilayah tertentu yang didasarkan pada jenis keluhan yang

spesifik, dengan demikian dapat memudahkan untuk melakukan analisis keluhan pelanggan sehingga pelayanan terhadap keluhan dapat ditingkatkan.

Penelitian ini hanya melakukan proses pengelompokan dengan teknik K-Medoids clustering. Penentuan nilai K sebagai jumlah klaster menggunakan perhitungan elbow serta dihitung menggunakan tools R studio.

2. METODE PENELITIAN

Metode penelitian berisi langkah-langkah yang sistematis yang dilakukan dalam proses pembuatan riset ini, yang terdiri dari proses pengumpulan data dan proses pengembangan perangkat lunak.

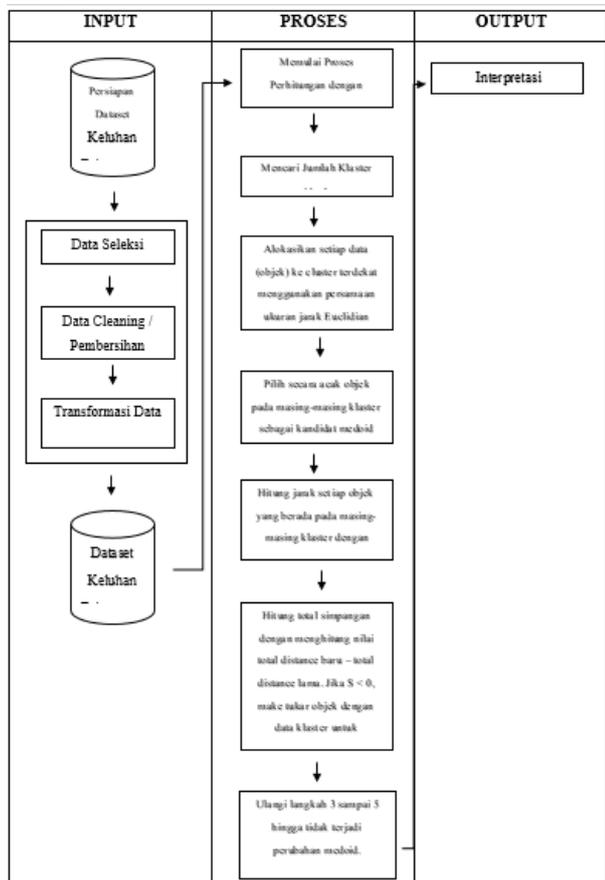
2.1. Perolehan Data

Prosedur yang digunakan dalam pengumpulan data untuk studi ini yaitu :

- a. Studi Pustaka dari penelitian sebelumnya yang memiliki hubungan dengan masalah yang diteliti.
- b. Observasi dengan cara mengunjungi Perumda Air Minum Tirta Raharja.
- c. Wawancara pada kepala PDAM Tirta Raharja Cimahi.

2.2. Pengembangan Perangkat Lunak

Proses pengembangan perangkat lunak terdiri dari beberapa tahap yaitu tahap pra proses, tahap pengelompokkan menggunakan K-Medoids, dan tahap pengujian. Perancangan sistem dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Metodologi Penelitian

a. Pra Proses

Proses penyaringan data yang tidak digunakan. Tahapan ini terdiri dari *data selection*, *data cleaning*, dan *transformation*.

b. Pengelompokan Menggunakan Metode K-Medoids

Proses pengelompokan data keluhan pelanggan menggunakan metode K-Medoids.

c. Pengujian

Proses pengujian pada setiap fungsi yang terdapat pada modul untuk mendeteksi apakah terdapat kesalahan pada sistem yang telah dibuat.

3. TINJAUAN PUSTAKA

Berikut penelitian terkait dengan data mining khususnya mengenai klasterisasi.

3.1. Data Mining

Data mining adalah serangkaian proses yang dilakukan yang bertujuan untuk menemukan pola dalam data. [5]. Secara lebih jelas *data mining* menggambarkan proses pencarian pola atau informasi dalam data terpilih dengan menggunakan metode ataupun tata cara tertentu [6].

3.2. Clustering

Clustering merupakan proses dari pembentukan segmen – segmen data dengan cara mengukur kemiripan sebuah data dengan data lainnya [7]. Tujuan dari klasterisasi adalah untuk meminimalisir variasi di dalam suatu *cluster* dan memaksimalkan variasi antar *cluster* [8].

3.3. Algoritma K-Medoids

Dalam [9] Yulison dan Gunawan (2015) mengungkapkan ” K-Medoids adalah sebuah algoritma yang merepresentasikan *cluster* yang dibentuk menggunakan medoids ”. Proses pembentukan *cluster* diawali dengan menentukan terlebih dahulu k objek dari *dataset* yang ada, proses tersebut dilakukan secara acak yang bertujuan untuk menjadikan objek tersebut sebagai medoid awal, setelah itu hitung setiap objek yang bukan medoid dengan k objek, total jarak terkecil setiap objek yang bukan medoid terhadap medoid akan masuk dalam *cluster* dimana medoid tersebut berada. Selanjutnya tentukan objek baru secara acak dan proses perhitungan dilakukan kembali. Apabila total jarak yang dihasilkan lebih kecil dari nilai setiap objek dengan medoid lama, maka objek baru tersebut dapat menjadi medoid baru. Iterasi terus dilakukan hingga tidak terjadi perubahan terhadap total jarak yang dihasilkan oleh medoid baru.

3.4. Metode Elbow

Metode elbow adalah sebuah cara yang bisa digunakan untuk mendapatkan informasi perihal menentukan jumlah *cluster* yang paling ideal untuk proses klasterisasi [10].

4. HASIL DAN PEMBAHASAN

Pada penelitian ini akan dibangun sistem segmentasi keluhan pelanggan di Perumda Air Minum Tirta Raharja Cimahi Menggunakan Metode k-Medoids.

4.1. Data Keluhan

Data keluhan pelanggan yang akan diolah nantinya sudah mengalami seleksi data untuk menghilangkan data yang tidak diperlukan seperti nama, tanggal dan lainnya sehingga hanya menyisakan wilayah (kelurahan/desa) dan jumlah keluhan yang masuk berdasarkan jenisnya. Data hasil seleksi kemudian diubah ke database berbentuk sql. Data hasil seleksi dapat dilihat pada Tabel 1 dan data yang sudah berbentuk sql dapat dilihat pada Gambar 2.

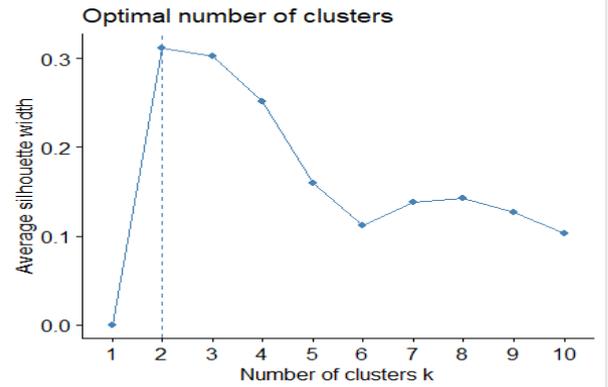
Tabel 1. Data Keluhan

Wilayah	A	B	C	D	E	F	G	H
Cigugur Girang	0	1	4	8	0	3	2	0
Cihanjuang Rahayu	0	0	6	9	0	6	2	0
Cihideung	2	3	5	4	2	3	5	1
Ciwaruga	0	1	5	8	0	2	3	1
Karyawangi	0	0	2	3	0	1	7	0
Sariwangi	0	0	3	10	0	3	5	0
Cipada	0	3	7	5	0	4	2	1
Tenjolaut	1	0	6	3	0	0	2	0
Batujajar Barat	0	0	5	4	0	5	5	0
Cangkorah	0	1	8	3	0	4	2	0
Pangauban	0	0	6	8	0	3	3	0
Budiharja	1	0	6	3	0	3	2	0

Gambar 2. Data Keluhan SQL

4.2. Proses Implementasi

Tahap ini merupakan proses pengelompokan terhadap data keluhan pelanggan yang sebelumnya telah disiapkan. Sebelum melakukan proses pengelompokan, mula - mula inisialisasikan pusat cluster sebanyak k (jumlah cluster). Diketahui jumlah k yang ideal menurut hasil pengujian metoda elbow adalah sebanyak 2, yang dapat dilihat pada Gambar 3.



Gambar 3. Jumlah k Ideal

Setelah diketahui jumlah k ideal maka proses pengelompokan dapat dilakukan. Berikut perhitungan yang dilakukan secara manual yang diimplementasikan pada data keluhan pelanggan di Perumda Air Minum Tirta Raharja Cimahi menggunakan metode k-Medoids.

Tabel 2. Atribut Data

No	Atribut	Inisialisasi
1	Pipa Distribusi	A
2	Pipa Dinas & Asesoris	B
3	Meter Air	C
4	Kontinuitas Pengaliran	D
5	Kualitas Air	E
6	Persil Bocor	F
7	Kesalahan Rekening	G
8	Lain - lain	H

1. Dipilih dua buah objek secara acak dari dataset yang ada untuk mewakili dua buah cluster. Hasil dari pembentukan cluster awal seperti pada Tabel 3 berikut :

Tabel 3. Objek Awal Sebagai Medoids

Nama Wilayah	A	B	C	D	E	F	G	H
Cigugur Girang	0	1	4	8	0	3	2	0
Budiharja	1	0	6	3	0	3	2	0

2. Berikutnya hitung jarak setiap objek non-medoids. Objek dengan jarak yang terdekat dengan masing-masing medoid akan bergabung dengan medoidnya. Tabel 4 berikut memperlihatkan pembentukan cluster yang dimulai dari iterasi ke-0 hingga selesai dan menghasilkan 2 buah cluster.

Tabel 4. Cluster Terbentuk

Wilayah	A	B	C	D	E	F	G	H	C1	C2	C
Cigugur Girang	0	1	4	8	0	3	2	0	0	5,57	1
Cihanjuang Rahayu	0	0	6	9	0	6	2	0	4,24	4,36	1
Cihideung	2	3	5	4	2	3	5	1	6,25	5,1	2
Ciwaruga	0	1	5	8	0	2	3	1	2	5,57	1
Karyawangi	0	0	2	3	0	1	7	0	7,68	6,78	2
Sariwangi	0	0	3	10	0	3	5	0	3,87	8,25	1
Cipada	0	3	7	5	0	4	2	1	4,9	4,12	2
Tenjolaut	1	0	6	3	0	0	2	0	4,32	3	2
Batujajar Barat	0	0	5	4	0	5	5	0	5,57	4	2
Cangkorah	0	1	8	3	0	4	2	0	6,48	2,65	2
Pangauban	0	0	6	8	0	3	3	0	2,45	5,2	1
Budiharja	1	0	6	3	0	3	2	0	5,57	0	2

Metode *Euclidean distance* digunakan untuk proses perhitungan jarak antara objek dengan setiap medoids, dengan persamaan sebagai berikut :

$$\sum_{i=1}^k (x_i - y_i)^2 \tag{1}$$

Euclidean Distance yaitu metode perhitungan jarak yang didasarkan pada ruang berdimensi terbatas yang bernilai riil [11].

Diketahui bahwa daerah yang terdapat pada *cluster* 1 adalah Cigugur Girang, Cihanjuang Rahayu, Ciwaruga, Sariwangi, dan Pangauban. Sementara daerah yang terdapat pada *cluster* 2 adalah Cihideung, Karyawangi, Cipada, Tenjolaut, Batujajar Barat, Cangkorah dan Budiharja seperti pada Tabel 5 dan Tabel 6.

Tabel 5. Data Wilayah C1

Nama Wilayah	A	B	C	D	E	F	G	H
Cigugur Girang	0	1	4	8	0	3	2	0
Cihanjuang Rahayu	0	0	6	9	0	6	2	0
Ciwaruga	0	1	6	8	0	2	3	1
Sariwangi	0	0	3	10	0	3	5	0
Pangauban	0	0	6	8	0	3	3	0
Rata - rata	0	0,02	0,235	0,42	0	0,17	0,15	0,01

Tabel 6. Data Wilayah C2

Nama Wilayah	A	B	C	D	E	F	G	H
Cihideung	2	3	5	4	2	3	5	1
Karyawangi	0	0	2	3	0	1	7	0
Cipada	0	3	7	5	0	4	2	1
Tenjolaut	1	0	6	3	0	0	2	0
Batujajar Barat	0	0	5	4	0	5	5	0
Cangkorah	0	1	8	3	0	4	2	0
Budiharja	1	0	6	3	0	3	2	0
Rata - rata	0,032	0,056	0,315	0,202	0,016	0,161	0,202	0,016

Untuk mengetahui keluhan mana yang lebih dominan maka hitung rata-rata keluhannya. Disini terlihat rata-rata keluhan yang sering muncul untuk C1 adalah jenis keluhan D dengan presentase 42% dari total keluhan yang ada pada *cluster* tersebut. Sementara untuk C2 adalah jenis keluhan C dengan presentase 31.5% dari total keluhan yang ada pada *cluster* tersebut.

5. KESIMPULAN

Hasil dari penelitian ini adalah sistem segmentasi keluhan pelanggan di perumda air minum tirta raharja cimahi menggunakan metode k-Medoids, yang berguna untuk mengetahui kecenderungan jenis keluhan di suatu wilayah tertentu. Keluhan yang sering muncul untuk C1 adalah jenis keluhan D dengan presentase 42% dari total keluhan yang ada pada *cluster* tersebut, sementara untuk C2 adalah jenis keluhan C dengan presentase 31.5% dari total keluhan yang ada pada *cluster* tersebut..

DAFTAR PUSTAKA

[1] Susi Indriyani and S. Mardiana, "Pengaruh Penanganan Keluhan (Complaint Handling) Terhadap Kepercayaan Dan Komitmen Mahasiswa," *J. Bus.*, vol. 2, no. 01, pp. 1–13, 2016.

[2] A. Nanang and D. Susanti, "Penerapan Data Mining Menggunakan Algoritma K-Means Untuk," *Sinergi*, pp.

93–102.

[3] D. F. Pramesti, Lahan, M. Tanzil Furqon, and C. Dewi, "Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 9, pp. 723–732, 2017, doi: 10.1109/EUMC.2008.4751704.

[4] D. Marlina, N. Lina, A. Fernando, and A. Ramadhan, "Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 4, no. 2, p. 64, 2018, doi: 10.24014/coreit.v4i2.4498.

[5] S. Saikin and K. Kusriani, "Model Data Mining Untuk Karakteristik Data Traveller Pada Perusahaan Tour and Travel," *J. Manaj. Inform. dan Sist. Inf.*, vol. 2, no. 2, p. 61, 2019, doi: 10.36595/misi.v2i2.105.

[6] Y. Wicaksono, "Segmentasi Pelanggan Bisnis Dengan Multi Kriteria Menggunakan K-Means," *Indones. J. Bus. Intell.*, vol. 1, no. 2, p. 45, 2019, doi: 10.21927/ijubi.v1i2.872.

[7] M. A. Syakur, B. K. Khotimah, E. M. S. Rochman, and B. D. Satoto, "Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 336, no. 1, 2018, doi: 10.1088/1757-899X/336/1/012017.

[8] F. A. Syam, "Implementasi Metode Klastering K-Means untuk Mengelompokan Hasil Evaluasi Mahasiswa," *J. Ilmu Komput. dan Bisnis*, vol. 8, no. 1, pp. 1857–1864, 2017, doi: 10.47927/jikb.v8i1.94.

[9] Y. H. Chrisnanto and G. Abdillah, "Penerapan Algoritma Partitioning Around Medoids (Pam) Clustering Untuk Melihat Gambaran Umum Kemampuan," *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2015, no. Sentika, pp. 444–448, 2015.

[10] N. Putu, E. Merliana, and A. J. Santoso, "Analisa Penentuan Jumlah Cluster Terbaik pada Metode K-Means," pp. 978–979.

[11] A. Kumari and H. Bhagat, "Compression Record Based Efficient k-Medoid Algorithm to Increase Scalability and Efficiency," vol. 2, no. 8, pp. 2398–2401, 2013.