

Klasterisasi Jemaat Penerima Bantuan Sosial Dengan K-Means Clustering

Mitha Febriana Sitanggang¹, Humuntal Rumapea², Indra M Sarkis³, Benget Rumahorbo⁴
^{1,2,3,4}Fakultas Ilmu Komputer, Universitas Methodist Indonesia

Info Artikel

Histori Artikel:

Received, Jan 19, 2022

Revised, Feb 19, 2022

Accepted, Maret 19, 2022

Keywords:

Klasterisasi,
K-Means,
Jemaat,
HKBP Resort Pardomuan Nauli

ABSTRAK

Proses klasterisasi jemaat yang layak menerima bantuan di Gereja HKBP Resort Pardomuan Nauli sedang mengalami masalah. Bervariasinya data jemaat membuat pihak gereja kurang objektif dalam memberikan bantuan sosial kepada jemaatnya. Untuk itu diperlukan sebuah metode pengklasifikasian jemaat yang penerima bantuan sosial untuk mengetahui dan mengelompokkan jemaat berdasarkan status ekonomi, sehingga mempermudah pihak gereja dalam menentukan jemaat mana saja yang layak menerima bantuan. Dalam pendekatan pengklasteran *K-Means*, pembagian kelompok jemaat dapat dilakukan berdasarkan status, usia, pekerjaan, tanggungan, pendidikan, dan status kepemilikan rumah (SKR). Dari penelitian yang dilakukan, terdapat 6 kelompok cluster yaitu C1= cluster kaya, C2 = cluster menengah atas, C3 = cluster menengah, C4 = cluster menengah bawah, C5 = cluster miskin, dan C6 = cluster sangat miskin, dengan tingkat akurasi 65,58% benar.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Penulis Koresponden:

Mitha Febriana Sitanggang,
Faculty of Computer Science, Universitas Methodist Indonesia, Medan,
Jl. Hang Tua No.8, Medan - Sumatera Utara.
Email: mithafeb@gmail.com

1. PENDAHULUAN

Gereja HKBP Resort Pardomuan Nauli merupakan organisasi yang mempunyai misi sosial untuk mengatasi masalah yang terjadi dalam sebuah masyarakat. Aksi sosial merupakan wujud gereja HKBP dalam membantu perekonomian masyarakat, maka dari itu gereja HKBP rutin memberikan bantuan sosial kepada jemaatnya. Kontribusi gereja ini tidak hanya sekedar untuk membantu jemaat miskin, tetapi juga berdampak bagi transformasi kehidupan yang lebih layak. Ada 7 pagaran yang dinaungi oleh gereja ini yaitu Gereja HKBP Bukit Maraja, Gereja HKBP serapu, Gereja HKBP Raya Timuran, Gereja HKBP Batu 8, Gereja HKBP Porsea Pardamean, Gereja HKBP Marihat Baris (Geraktani), dan Gereja HKBP Pondok Baru (Marihat).

Aksi sosial yang dilakukan oleh Gereja HKBP Resort Pardomuan Nauli ini merupakan kegiatan rutin yang dilaksanakan oleh Bidang Diakonia yang dilakukan 2 kali dalam setahun yaitu pada Bulan Mei dan Bulan Desember ketika Natal. Bantuan sosial yang diberikan Gereja bersumber dari jemaat untuk jemaat, salah satu caranya dengan mengadakan pagelaran koor per-3 bulan di setiap tahun yang merupakan program kerja dari Bidang Koinonia. Sumbangan yang diberikan jemaat saat pagelaran koor nantinya akan digunakan untuk aksi sosial. Bantuan (saguhati) yang diberikan oleh gereja juga tidak sama bentuknya, sesuai dengan kelompok-kelompok yang telah ditentukan. Aksi sosial yang dilakukan oleh gereja HKBP

seperti memberi uang dan sembako berupa beras, sepanan telur, dan buah-buahan kepada jemaat yang miskin dan sangat miskin.

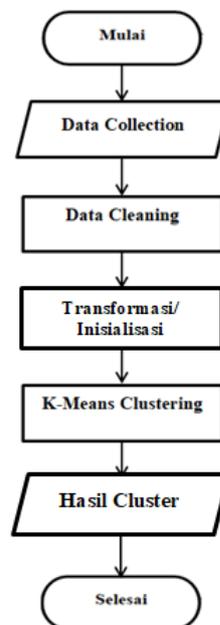
Jumlah dan status ekonomi jemaat di gereja HKBP Resort Pardomuan Nauli mengalami perubahan setiap tahunnya. Perubahan ini diakibatkan oleh perubahan kesenjangan ekonomi jemaat dan keluar masuknya (perpindahan) jemaat yang membuat pihak gereja harus selalu *update* mengenai perubahan-perubahan yang terjadi. Dengan keterbatasan dana bantuan yang dimiliki Gereja, membuat pihak Gereja harus benar-benar objektif dalam memilih jemaat yang layak menerima bantuan. Selain itu, Gereja HKBP Resort Pardomuan Nauli juga belum memiliki sistem untuk mengelompokkan jemaat kedalam kelompok-kelompok yang layak menerima bantuan, sehingga terkadang bantuan yang diberikan tidak pada orang yang tepat (masih kurang objektif). Dari permasalahan ini, penulis tertarik melakukan penelitian menggunakan metode *K-Means Clustering* dalam pengelompokan pembagian bantuan sosial kepada jemaat di Gereja HKBP Resort Pardomuan Nauli.

Algoritma K-Means merupakan algoritma klasterisasi yang mengelompokkan data berdasarkan titik pusat klaster (*centroid*) terdekat dengan data. Tujuan dari K-Means adalah mengelompokkan data dengan memaksimalkan kemiripan data dalam satu klaster dan meminimalkan kemiripan data antar klaster (Darmi & Setiawan, 2016). Ukuran kemiripan yang digunakan dalam klaster adalah fungsi jarak. Sehingga pemaksimalan kemiripan data didapatkan berdasarkan jarak terpendek antara data terhadap titik centroid. Dengan data yang sudah dikelompokkan menggunakan algoritma *K-Means* diharapkan dapat mempermudah Gereja HKBP Resort Pardomuan Nauli dalam menentukan jemaat yang layak menerima bantuan.

2. METODE PENELITIAN

Adapun proses atau langkah-langkah yang digunakan dalam penelitian ini dapat dilihat pada gambar berikut.

2.1 Framework Penelitian



Gambar 1. Framework Penelitian

Pada gambar 1 proses *clustering* di mulai dari tahap pengumpulan, pada tahap ini akan dilakukan pengumpulan data langsung dari Gereja HKBP Reosort Pardomuan Nauli. Pengumpulan data di lakukan dengan proses wawancara terhadap *stake holder* (Pendeta, Majelis, ataupun Sintua di Gereja HKBP Resort Pardomuan Nauli). Input data mentah yang berupa data jemaat yang berektensi .xlsx akan mengalami representasi dari data. Tahap kedua yaitu tahap data pembersihan (*cleaning*) untuk membuang data yang tidak konsisten. Apabila sudah selesai maka didapatkan data yang sudah diproses atau data matang siap untuk ke tahap selanjutnya. Tahap ke tiga yaitu tranformasi data, data yang berjenis alfabet seperti status, jenis pekerjaan, dll harus dilakukan proses inisialisasi data terlebih dahulu ke dalam bentuk

angka/numerikal. Tahap ke empat yaitu *K-Means Clustering* disini tahap *K-Means clustering* adalah mengelompokkan data yang sudah ada ke dalam enam kelompok yaitu kaya, menengah atas, menengah, menengah bawah, miskin, dan sangat miskin.

2.2 Metode K-Means Clustering

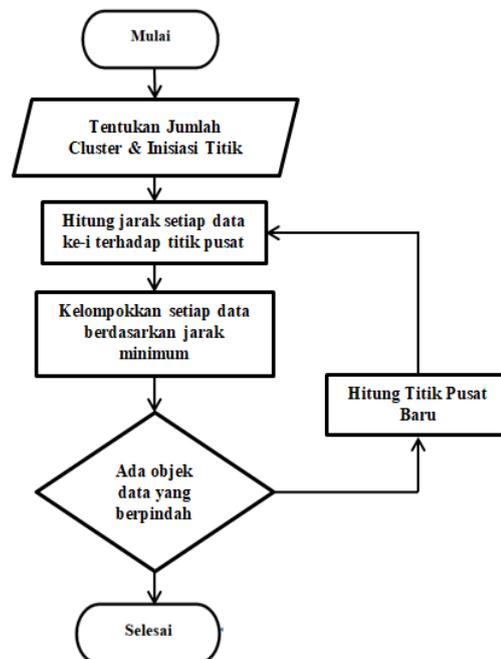
Proses *clustering* dengan menggunakan algoritma *K-Means* memiliki langkah-langkah sebagai berikut:

- Inisialisasi : menentukan K sebagai jumlah *cluster* yang diinginkan dan metrik ketidakmiripan (jarak) yang diinginkan. Jika perlu, tetapkan ambang batas perubahan fungsi objektif dan ambang batas perubahan *centroid*.
- Pilih K data baru set data X sebagai *centroid*.
- Alokasikan semua data ke *centroid* terdekat dengan metrik jarak yang sudah ditetapkan (memperbaharui ID setiap data).
- Hitung kembali *centroid* C berdasarkan data yang mengikuti *cluster* masing-masing.
- Ulangi langkah tiga dan empat hingga kondisi konvergen tercapai, yaitu (a) perubahan fungsi objektif sudah dibawah ambang batas yang diinginkan; atau (b) tidak ada data yang berpindah *cluster* ; atau (c) perubahan posisi *centroid* sudah dibawah ambang batas yang ditetapkan.

3. HASIL DAN PEMBAHASAN

3.1. Proses K-Means Clustering

Pemodelan K-Means dimodelkan dalam algoritma flowchart, seperti gambar 2. Flowchart Algoritma K-Means.



Gambar 2. Flowchart Algoritma K-Means

Untuk dapat melakukan pengelompokan data menjadi beberapa *cluster* perlu dilakukan beberapa langkah, yaitu:

- Tentukan jumlah *cluster* yang diinginkan. Dalam penelitian ini data-data yang ada akan dikelompokkan menjadi 6 *cluster* yaitu C1 = Kaya, C2 = Menengah Atas, C3 = Menengah, C4 = Menengah Bawah, C5 = Miskin, dan C6 = Sangat Miskin.
- Tentukan titik pusat awal dari setiap *cluster*. Dalam penelitian ini titik pusat awal ditentukan secara random dan didapat titik pusat dari setiap *cluster* dapat dilihat pada tabel 1 dan sampel data yang digunakan dapat dilihat pada tabel 2.

Tabel 1. Penentuan Titik Pusat Cluster

No	Titik Pusat Awal	Status	Usia	Pekerjaan	Tanggung	Pendidikan	SKR
1	C1	1	1	5	1	1	1
2	C2	1	3	5	0	4	2
3	C3	2	2	5	3	4	2
4	C4	2	2	5	0	4	1
5	C5	3	1	8	0	2	8
6	C6	1	2	5	4	4	2

Tabel 2. Sampel Data Jemaat Pada Gereja HKBP Resort Pardomuan Nauli

NO	Nama Kepala Keluarga	Status	Usia	Pekerjaan	Tanggung	Pendidikan	SKR
1	OP. MUTIARA BR SIANTURI	2	2	1	0	2	1
2	OP. DEPIN BR LUMBANGAOL	2	3	1	0	3	1
3	OP. NATAN BR SIAGIAN	2	3	1	0	2	1
4	OP. JANET SITOMPUL	1	3	3	1	2	2
5	OP. OKKTA TAMBUNAN	1	2	4	2	0	1
6	OP. PRILIA SIRAIT	1	3	1	0	1	2
7	OP. RICA SIAHAAN	1	2	0	1	1	1
8	AMA VAREL SILITONGA	1	1	2	2	1	1
9	AMA INGGRID SILALAH	1	1	3	1	2	1
10	OP. RISDO BR TAMBUNAN	2	2	5	3	5	1
11	OP. INGGRIT SILALAH	1	2	2	2	0	1
12	St.TB.HUTAGAOL	1	3	1	1	1	1
13	AMA ROHANA HUTABARAT	1	1	0	3	1	1
14	AMA PEDRO NAPITUPULU	1	2	3	5	2	1
15	AMA MARUDUT HASIBUAN	1	2	5	4	5	2
16	St. J. MANIK	1	2	0	2	0	1
17	NAI ROY BR SIHOMBING	2	2	7	3	2	2
18	OP. VIONA PASARIBU	1	2	0	2	1	1
19	OP. TIKA BR SIPAHUTAR	2	2	4	0	3	1
20	AMA DANIEL NABABAN	1	1	0	2	0	1

- c) Setelah diketahui nilai k dan pusat cluster awal, selanjutnya mengukur jarak antara pusat cluster menggunakan *euclidian distance*, kemudian akan didapatkan matriks jarak yaitu C1, C2, C3, C4, C5, dan C6 sebagai berikut:

Rumus *euclidian distance*:

$$d(x,y) = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Perhitungan jarak data pertama pada pusat cluster pertama:

$$d_{11} = \sqrt{(2-1)^2 + (2-1)^2 + (1-5)^2 + (0-1)^2 + (2-1)^2 + (1-1)^2} = 4,3589$$

Perhitungan jarak data pertama pada pusat cluster kedua:

$$d_{12} = \sqrt{(2-1)^2 + (2-3)^2 + (1-5)^2 + (0-0)^2 + (2-4)^2 + (1-2)^2} = 4,358899$$

Perhitungan jarak data pertama pada pusat cluster ketiga:

$$d_{13} = \sqrt{(2-2)^2 + (2-2)^2 + (1-5)^2 + (0-3)^2 + (2-4)^2 + (1-2)^2} = 7,28011$$

Perhitungan jarak data pertama pada pusat cluster keempat:

$$d_{14} = \sqrt{(2-2)^2 + (2-2)^2 + (1-5)^2 + (0-0)^2 + (2-4)^2 + (1-1)^2} = 4$$

Perhitungan jarak data pertama pada pusat cluster kelima:

$$d_{15} = \sqrt{(2-3)^2 + (2-1)^2 + (1-8)^2 + (0-0)^2 + (2-2)^2 + (1-8)^2} = 10$$

Perhitungan jarak data pertama pada pusat cluster keenam:

$$d_{16} = \sqrt{(2-1)^2 + (2-2)^2 + (1-5)^2 + (0-4)^2 + (2-4)^2 + (1-2)^2} = 9,0554$$

dst.

Tabel 1 Hasil Perhitungan Jarak Cluster

No	C1	C2	C3	C4	C5	C6	Terdekat
1	4,3589	4,358899	7,28011	4	10	9,0553851	4
2	5	4,242641	5,196152	4,12311	10,1489	5,9160798	4,12311
3	4,69042	4,242641	7,348469	4,12311	10,1489	9,1104336	4,12311
4	3	2,828427	4,690416	3,31662	8,30662	6,4031242	2,82843
5	1,73205	8,185353	4,358899	8,12404	9,27362	8,1240384	1,73205
6	4,58258	4	9,949874	4,3589	9,64365	12,688578	4
7	5,09902	6	7,937254	5,91608	10,9087	10,34408	5,09902
8	3	7,071068	4,582576	6,85565	9,64365	6,8556546	3
9	2	3,605551	4,795832	3,16228	8,83176	6,4807407	2
10	8,12404	3,464102	1	3	11,8743	1,7320508	1
11	3,31662	8,660254	5,196152	8,60233	10,2956	8,6023253	3,31662
12	4,47214	5,09902	7,416198	5,19615	10,3441	9,9498744	4,47214
13	5	10,53565	5,291503	10,3923	11,225	6	5
14	4,58258	10,29563	4,690416	10,247	8,88819	3	3
15	12,083	4,123106	1,414214	4,24264	13,9284	0	0
16	5,19615	9,539392	6,557439	9,48683	11,5758	9,486833	5,19615
17	3,31662	6,480741	2	6,40312	6,245	3	2
18	5,09902	7,937254	6	7,87401	11,0454	7,8740079	5,09902
19	2,64575	2	3,316625	1	8,18535	4,3588989	1
20	5,09902	9,69536	6,63325	9,53939	11,5326	9,539392	5,09902

- d) Langkah selanjutnya hasil perhitungan akan dilakukan perbandingan dipilih jarak terdekat antara data dengan pusat cluster, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusatcluster terdekat. Berikut ini akan ditampilkan data matriks pengelompokan group, nilai 1 berarti data tersebut berada dalam *group* (kelompok data yang sama).

e) Tabel 4. Pengelompokan Group Iterasi

No	C1	C2	C3	C4	C5	C6
1				1		
2				1		
3				1		
4		1				
5	1					
6		1				
7	1					
8	1					

9	1					
10		1				
11	1					
12	1					
13	1					
14					1	
15					1	
16	1					
17		1				
18	1					
19				1		
20	1					

Setelah diketahui anggota tiap-tiap cluster kemudian pusat cluster baru dihitung berdasarkan data anggota tiap-tiap cluster sesuai dengan rumus pusat anggota cluster untuk menguji apakah terjadi perpindahan kelompok pada data tersebut. Jika ada terjadi perpindahan kelompok pada data, maka harus dilakukan penentuan kelompok kembali pada iterasi selanjutnya dengan nilai titik pusat cluster yang harus diupdate ulang, tetapi jika tidak terjadi perpindahan kelompok pada setiap data maka data telah mencapai stabil dan konvergen.

3.2. Hasil Akurasi

Pengujian hasil dilakukan dengan membandingkan pengelompokan yang dilakukan oleh algoritma K-Means oleh sistem dengan pengelompokan yang dilakukan oleh pihak Gereja HKBP Resort Prdomuan Nauli. Dengan membandingkan tabel pengelompokan jemaat penerima bantuan dari Gereja HKBP Resort Pardomuan Nauli (data latih) dan pengelompokan dengan algoritma K-Means didapatkan hasil sebagai berikut :

- Terdapat 141 jemaat dengan hasil pengelompokan yang sama
- 74 jemaat lainnya terdapat perbedaan dalam penentuan kelompok atau cluster nya.

Precision

	c1	c2	c3	c4	c5	c6
c1	18	0	0	0	0	0
c2	2	10	0	0	0	0
c3	0	0	8	1	0	1
c4	2	6	11	41	1	0
c5	3	10	23	14	59	0
c6	0	0	0	0	0	5

actual

Gambar 3. Confusion Matrix

Formula Akurasi :

Akurasi = TP / Total Dataset

Akurasi = $18 + 10 + 8 + 41 + 59 + 5 / 215 \times 100 = 65,58\%$

Dari hasil tersebut dapat disimpulkan bahwa kebenaran algoritma K-Means dalam mengklasterisasi jemaat memiliki nilai akurasi sebesar 65,58% (141/215).

4. KESIMPULAN

Adapun kesimpulan yang dapat di ambil pada penelitian ini adalah Algoritma *K-Means Clustering* ini dapat digunakan untuk mengklasterisasi jemaat penerima bantuan sosial di Gereja HKBP Resort Pardomuan Nauli. Dengan membandingkan 215 data latih pengelompokan dari Gereja HKBP Resort Pardomuan Nauli dan algoritma K-Means didapat 141 data yang memiliki kelompok yang sama. Dari angka ini dapat disimpulkan bahwa algoritma K-Means untuk mengklasterisasi jemaat penerima bantuan sosial di Gereja HKBP Resort Pardomuan Nauli memiliki nilai akurasi 65,58% benar. Nilai ini bisa berubah seiring dengan penambahan data latih.

REFERENSI (11 PT)

- [1] Darmi, Y., & Setiawan, A. (2016). Penerapan Metode Clustering K-Means Dalam. *Y. Darmi, A. Setiawan, 12(2)*, 148–157.
- [2] F.N.R. Fauzan, J. Aziz, B.D. Setiawan, dan I. Arwani, “Implementasi Algoritma K-Means untuk Klasterisasi Kinerja Akademik Mahasiswa”, *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2. No. 6, 2018.
- [3] Gustientiedina, G., Adiya, M. H., & Desnelita, Y. (2019). Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan. *Jurnal Nasional Teknologi dan Sistem Informasi*, 5(1), 17–24. <https://doi.org/10.25077/teknosi.v5i1.2019.17-24>
- [4] Hasanah, N., Ugiarto, M., & Puspitasari, N. (2017). Sistem pengelompokan curah hujan menggunakan metode k-means di wilayah kalimantan timur. *Prosiding Seminar Nasional Ilmu Komputer Dan Teknologi Informasi*, 2(2), 122–126.
- [5] Helilintar, R., & Farida, I. N. (2018). Penerapan Algoritma K-Means Clustering Untuk Prediksi Prestasi Nilai Akademik Mahasiswa. *Jurnal Sains dan Informatika*, 4(2), 80–87. <https://doi.org/10.34128/jsi.v4i2.140>
- [6] Isa, I. G. T., & Hartawan, G. P. (2017). Perancangan Aplikasi Koperasi Simpan Pinjam Berbasis Web (Studi. *Jurnal Ilmiah Ilmu Ekonomi*, 5(10), 139–151.
- [7] Maulana, A., & Fajrin, A. A. (2018). Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen Dengan Algoritma Fp-Growth Pada Data Transaksi Penjualan Spare Part Motor. *Klik - Kumpulan Jurnal Ilmu Komputer*, 5(1), 27. <https://doi.org/10.20527/klik.v5i1.100>
- [8] Muttaqin, M. R., & Defriani, M. (2020). Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa. *ILKOM Jurnal Ilmiah*. <https://doi.org/10.33096/ilkom.v12i2.542.121-129>
- [9] Noviyanto, N. (2020). Penerapan Data Mining dalam Mengelompokkan Jumlah Kematian Penderita COVID-19 Berdasarkan Negara di Benua Asia. *Paradigma - Jurnal Komputer dan Informatika*, 22(2), 183–188. <https://doi.org/10.31294/p.v22i2.8808>
- [10] Parlina, I., Windarto, A. P., Wanto, A., & Lubis, M. R. (2018). Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Asessment Center. *Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Asessment Center Untuk Clustering Program Sdp*, 3(1), 87–93.
- [11] S. Sulistyowati, B.E. Ketherin, A.A. Arifiyanti, dan A, Sodik. Analisa Segmentasi Konsumen Menggunakan Algoritma K-Means Clustering, *Prosiding Seminar Nasional Sains dan Teknologi Terapan VI (SNTEKPAN VI)*, 2018.
- [12] Widodo dan D.Wahyuni, "Implementasi Algoritma K-Means Clustering Untuk Mengetahui Bidang Skripsi Mahasiswa Multimedia Pendidikan Teknik Informatika Dan Komputer Universitas Negeri Jakarta", *JURNAL Pendidikan Teknik Informatika dan Komputer*, Vol 1 No 2, 2017.