

Penerapan Metode K-Medoid pada Pengelompokan Rumah Tangga Dalam Perlakuan Memilah Sampah Menurut Provinsi

Desi Asima Silitonga, Agus Perdana Windarto, Dedy Hartama, Sumarno

STIKOM Tunas Bangsa, Pematangsiantar Indonesia

Email: ¹silitongadesi03@gmail.com, ²agus.perdana@amiktunasbangsa.ac.id

Abstrak

Indonesia merupakan negara penyumbang sampah terbesar kedua didunia. Banyak masyarakat disetiap wilayah yang ada di indonesia belum dapat menjaga kebersihan lingkungan, sehingga masalah sampah saat ini belum bisa ditangani dengan baik karena kurangnya kesadaran dan pemahaman manusia atau masyarakat dalam memilah dan mengelolah sampah. Sumber data yang digunakan dari situs Badan Pusat Statistik Indonesia yaitu data Persentase Rumah Tangga Menurut Provinsi dan Perlakuan, 2013-2014. Penelitian ini bertujuan untuk mengelompokkan perilaku memilah sampah dengan menggunakan algoritma datamining dengan K-Medoid. Data yang diolah menjadi 2 cluster yaitu cluster tingkat memilah sampah rendah (C1) dan cluster tingkat memilah sampah tinggi (C2). Dimana hasil penelitian ini menyimpulkan dari 33 provinsi di indonesia bahwa cluster tingkat perilaku memilah sampah rendah (C1) diperoleh 22 provinsi yaitu Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kep. Bangka Belitung, Kep. Riau, DKI Jakarta, Jawa Timur, Banten, NTB, Kalimantan Barat, Kalimantan Selatan, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua, dan 12 provinsi lainnya termasuk dalam cluster tingkat memilah sampah tinggi (C2).

Kata Kunci: Memilah Sampah, Datamining, K-Medoid, Clustering, Rumah Tangga

1. PENDAHULUAN

Indonesia merupakan negara penyumbang sampah terbesar kedua didunia. Undang-Undang Republik Indonesia Nomor 18 Tahun 2008 tentang sampah mengatakan bahwasampah telah menjadi permasalahan nasional sehingga pengelolannya perlu dilakukan secara komprehensif dan terpadu dari hulu ke hilir agar memberikan manfaat secara ekonomi, sehat bagi masyarakat, dan aman bagi lingkungan, serta dapat mengubah perilaku masyarakat (UU RI No 18, 2008)[1]. Banyak wilayah yang ada di indonesia belum dapat menjaga kebersihan lingkungan, sehingga masalah sampah saat ini belum bisa ditangani dengan baik sehingga banyak kerugian yang dirasakan mulai dari kesehatan manusia atau makhluk lainnya dan dapat mencemari lingkungan. Banyaknya timbunan sampah disebabkan oleh kurangnya kesadaran dan pemahaman manusia atau masyarakat dalam mengelolah sampah. Pengelolaan sampah dapat dilakukan dengan cara memilah atau memisahkan sampah berdasarkan jenis sampah yaitu sampah organik dan sampah non organik. Maka jika ada proses pemilahan sampah yang dilakukan oleh masyarakat tersebut setidaknya dapat mengurangi timbunan sampah yang ada di wilayah tersebut dan dapat memudahkan petugas untuk pemusnahan dan pengolahan sampah. Saat ini masyarakat masih butuh arahan dari Dinas Kebersihan dalam upaya pemilahan sampah. Penelitian ini fokus pada wilayah indonesia yang masih kurang melakukan tindakan dalam memilah sampah. Objek penelitian ini menitik beratkan perlakuan rumah tangga dalam memilah sampah menurut provinsi dengan menggunakan datamining metode clustering.

Berdasarkan permasalahan diatas, penulis mengangkat topik perlakuan dalam memilah sampah dengan data yang diambil berdasarkan provinsi dimana proses metode yang digunakan adalah clustering. Analisis kluster merupakan suatu teknik multivariat dengan tujuan utama mengelompokkan objek-objek berdasarkan karakteristik yang dimiliki[2]. Sekarang ini analisis kluster telah banyak diaplikasikan dalam berbagai bidang yang ditulis dalam berbagai penelitian dan jurnal[3]. Proses pengelompokkan dapat dilakukan dengan mengimplementasikan metode *clustering*[4], salah satunya yaitu metode *K-Medoids*[5]. Algoritma k-medoids dapat membantu dalam pengelompokan provinsi mana saja yang masih kurang dalam memilah sampah. Strategi dasar dari algoritma clustering K-Medoids adalah untuk menemukan k cluster dalam n objek dengan pertama kali secara arbitrarly menemukan wakil dari objek (medoid) untuk tiap-tiap cluster[6]. Data yang digunakan pada penelitian ini diambil melalui situs <https://www.bps.go.id> yang dihasilkan oleh Direktorat Jenderal Bea dan Cukai. Hasil pengklasteran bisa digunakan pemerintah dengan tujuan untuk membuat kebijakan disetiap wilayah atau daerah yang ada di indonesia.

2. METODOLOGI PENELITIAN

2.1 Pengumpulan Data

Data yang dibutuhkan dalam penelitian ini adalah data perlakuan Rumah Tangga/masyarakat dalam memilah sampah berdasarkan provinsi yang diperoleh dari sebuah situs resmi <https://www.bps.go.id>. Data yang dikumpulkan berupa data kelakuan masyarakat atas sampah yang tidak dipilah dari setiap provinsi yang ada di indonesia dengan jumlah 32 provinsi dari tahun 2013 hingga 2014. Variabel yang digunakan: Sampah tidak dipilah tahun 2013 (x), dan Sampah tidak dipilah tahun 2014.

2.2 Pengolahan Data

Pada pengolahan data ini penulis melakukan pengolahan data awal yang merupakan tahap untuk mempersiapkan data yang telah diperoleh sebelumnya dan akan digunakan pada tahap selanjutnya. Data yang sudah dikumpulkan akan diolah menjadi 2 cluster dengan beberapa tahap dan proses perhitungannya menggunakan bantuan aplikasi rapidminer.

2.3 Clustering

Clustering atau klusterisasi adalah salah satu alat bantu pada data mining yang bertujuan mengelompokkan objek-objek ke dalam cluster - cluster. Cluster adalah sekelompok atau sekumpulan objek - objek data yang similar satu sama lain dalam cluster yang sama dan disimilar terhadap objek-objek yang berbeda cluster. Objek akan dikelompokkan ke dalam satu atau lebih cluster sehingga objek - objek yang berada dalam satu cluster akan mempunyai kesamaan yang tinggi antara satu dengan lainnya.[7]

2.4 Algoritma K-Medoid

Metode K-Medoids merupakan bagian dari partitioning clustering. Metode K-Medoids cukup efisien untuk dataset yang kecil. Algoritma K-Medoids hadir untuk mengatasi kelemahan Algoritma K-Means[8] yang sensitif terhadap outlier karena suatu objek dengan suatu nilai yang besar mungkin secara substansial menyimpang dari distribusi data[6]. Langkah awal K-Medoids adalah mencari titik yang paling representatif (medoids) dalam sebuah dataset dengan menghitung jarak dalam kelompok dari semua kemungkinan kombinasi dari medoids sehingga jarak antar titik dalam suatu cluster kecil sedangkan jarak titik antar cluster besar[9]. Algoritma K-Medoids memiliki kelebihan untuk mengatasi kelemahan pada algoritma K-Means yang sensitive terhadap noise dan outlier, dimana objek dengan nilai yang besar yang memungkinkan menyimpang pada dari distribusi data[5].

Langkah- langkah algoritma K-Medoids:

- a. Inisialisasi pusat cluster sebanyak k (jumlah cluster)

Dalam penelitian ini penulis menggunakan 2 cluster yaitu: (C1) dan (C2)

- b. Alokasikan setiap data (objek) ke cluster terdekat menggunakan persamaan ukuran jarak Euclidian Distance dengan persamaan:

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

- c. Pilih secara acak objek pada masing-masing cluster sebagai kandidat medoid baru.
- d. Hitung jarak setiap objek yang berada pada masing-masing cluster dengan kandidat medoid baru.
- e. Hitung total simpangan (S) dengan menghitung nilai total distance baru – total distance lama. Jika $S < 0$, maka tukar objek dengan data cluster untuk membentuk sekumpulan k objek baru sebagai medoid.
- f. Ulangi langkah 3 sampai 5 hingga tidak terjadi perubahan medoid, sehingga didapatkan cluster beserta anggota cluster masing-masing[5].

Adapun skema alur tahapan penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Alur Tahapan Penelitian

3. ANALISA DAN PEMBAHASAN

Pada pengclusteran ini, data yang sudah dikumpulkan akan dihitung terlebih dahulu berdasarkan data perlakuan rumah tangga dalam memilah sampah dari 32 provinsi yang ada di Indonesia yang ditunjukkan pada Tabel 1.

Tabel 1. Data Awal Perlakuan Rumah Tangga terhadap sampah yang tidak dipilah

Provinsi	2013	2014
	Sampah Tidak dipilah	Sampah Tidak dipilah
Aceh	81,21	82,77
Sumatera Utara	80,39	84,98
Sumatera Barat	82,53	86,95
Riau	79,13	83,96
Jambi	83,90	88,85

Sumatera Selatan	76,82	87,70
Bengkulu	81,10	89,24
Lampung	83,71	84,69
Kepulauan Bangka Belitung	76,68	86,42
Kepulauan Riau	79,99	84,23
DKI Jakarta	85,77	88,65
Jawa Barat	69,48	77,36
Jawa Tengah	72,59	77,96
DI Yogyakarta	68,74	72,40
Jawa Timur	80,07	84,51
Banten	81,58	82,69
Bali	68,83	74,33
Nusa Tenggara Barat	82,17	88,29
Nusa Tenggara Timur	70,37	72,18
Kalimantan Barat	84,20	84,98
Kalimantan Tengah	76,16	75,87
Kalimantan Selatan	79,89	83,33
Kalimantan Timur	70,97	80,43
Sulawesi Utara	65,05	71,51
Sulawesi Tengah	70,05	79,88
Sulawesi Selatan	71,42	68,11
Sulawesi Tenggara	73,22	77,75
Gorontalo	77,75	85,62
Sulawesi Barat	79,48	85,78
Maluku	84,41	85,34
Maluku Utara	83,41	91,82
Papua Barat	72,02	74,10
Papua	83,02	81,90

Sumber : Badan Pusat Statistik , url : <https://www.bps.go.id>

3.1 Centroid Data

Dalam penerapan algoritma K-Medoid nilai titik tengah atau centroid dapat ditentukan secara random atau acak, dari data yang didapat dengan ketentuan bahwa clusterisasi yang diinginkan adalah 2, Penentuan cluster dibagi atas dua bagian yakni cluster tingkat memilah sampah tinggi (C1) dan cluster tingkat memilah sampah rendah (C2), maka nilai titik tengah atau centroid juga terdapat 2 titik. Penentuan titik cluster ini dilakukan dengan mengambil nilai secara acak. Data awal akan diolah menggunakan rumus Euclidian Distance dengan centroid awal yang telah ditentukan. Berikut adalah centroid awal yang digunakan:

Tabel 2. Data Centroid awal

Jawa Timur	80,07	84,51
Bali	68,83	74,33

Ada pun hasil perhitungannya sebagai berikut:

Tabel 3. Data Setelah Diolah

Provinsi	2013	2014	C1	C2	Jarak Pendek
	Sampah Tidak dipilah	Sampah Tidak dipilah			
Aceh	81,21	82,77	2,080192299	14,98325732	2,080192299
Sumatera Utara	80,39	84,98	0,568594759	15,71801832	0,568594759
Sumatera Barat	82,53	86,95	3,464852089	18,626712	3,464852089
Riau	79,13	83,96	1,089082182	14,10059928	1,089082182
Jambi	83,90	88,85	5,788307179	20,92690374	5,788307179
Sumatera Selatan	76,82	87,70	4,553965305	15,57552567	4,553965305
Bengkulu	81,10	89,24	4,840847033	19,30960901	4,840847033
Lampung	83,71	84,69	3,644447832	18,13129891	3,644447832
Kepulauan Bangka Belitung	76,68	86,42	3,891040992	14,41494363	3,891040992
Kepulauan Riau	79,99	84,23	0,291204396	14,91829749	0,291204396
DKI Jakarta	85,77	88,65	7,04482789	22,18165909	7,04482789
Jawa Barat	69,48	77,36	12,77773845	3,098935301	3,098935301
Jawa Tengah	72,59	77,96	9,94247957	5,226327583	5,226327583



DI Yogyakarta	68,74	72,40	16,58375711	1,932097306	1,932097306
Jawa Timur	80,07	84,51	0	15,16476179	0
Banten	81,58	82,69	2,364846718	15,2463799	2,364846718
Bali	68,83	74,33	15,16476179	0	0
Nusa Tenggara Barat	82,17	88,29	4,324164659	19,30899272	4,324164659
Nusa Tenggara Timur	70,37	72,18	15,68817708	2,644636081	2,644636081
Kalimantan Barat	84,20	84,98	4,156657311	18,6991818	4,156657311
Kalimantan Tengah	76,16	75,87	9,483548914	7,490026702	7,490026702
Kalimantan Selatan	79,89	83,33	1,193649865	14,25915846	1,193649865
Kalimantan Timur	70,97	80,43	9,972782962	6,464487605	6,464487605
Sulawesi Utara	65,05	71,51	19,86455134	4,716015267	4,716015267
Sulawesi Tengah	70,05	79,88	11,03799348	5,682508249	5,682508249
Sulawesi Selatan	71,42	68,11	18,54137266	6,737692483	6,737692483
Sulawesi Tenggara	73,22	77,75	9,623933707	5,56493486	5,56493486
Gorontalo	77,75	85,62	2,571867026	14,38855448	2,571867026
Sulawesi Barat	79,48	85,78	1,400357097	15,63729516	1,400357097
Maluku	84,41	85,34	4,418653641	19,07764398	4,418653641
Maluku Utara	83,41	91,82	8,036896167	22,77007905	8,036896167
Papua Barat	72,02	74,10	13,15943008	3,198280788	3,198280788
Papua	83,02	81,90	3,938857702	16,08294127	3,938857702

3.2 Clustering

Dengan menggunakan centroid tersebut maka dapat dicluster data yang telah didapat menjadi 2 cluster. Proses cluster dengan mengambil jarak terdekat dari setiap data yang diolah. Dari data awal didapatkan pengelompokan pada iterasi 1 untuk 2 cluster tersebut. Pada iterasi awal pengelompokan yaitu 2 cluster yang ditunjukkan pada Tabel 4.

Tabel 4. Hasil Cluster Iterasi Pertama

Klaster 1	Klaster 2
Aceh	Jawa Barat
Sumatera Utara	Jawa Tengah
Sumatera Barat	DI Yogyakarta
Riau	Bali
Jambi	Nusa Tenggara Timur
Sumatera Selatan	Kalimantan Tengah
Bengkulu	Kalimantan Timur
Lampung	Sulawesi Utara
Kepulauan Bangka Belitung	Sulawesi Tengah
Kepulauan Riau	Sulawesi Selatan
DKI Jakarta	Sulawesi Tenggara
Jawa Timur	Papua Barat
Banten	
Nusa Tenggara Barat	
Kalimantan Barat	
Kalimantan Selatan	
Gorontalo	
Sulawesi Barat	
Maluku	
Maluku Utara	
Papua	

Proses Clustering berhenti jika Simpangan > 0, tetapi jika Simpangan < 0 maka tukar objek dengan data cluster untuk membentuk sekumpulan k objek baru sebagai medoid. Proses perhitungan simpangan dengan cara menjumlahkan nilai total distance baru – total distance lama yang beisikan jarak setiap objek yang berada pada masing-masing cluster dengan kandidat medoid baru. Dengan jumlah simpangan yang ditunjukkan pada Tabel 5.

Tabel 5. Jumlah Distance Iterasi pertama

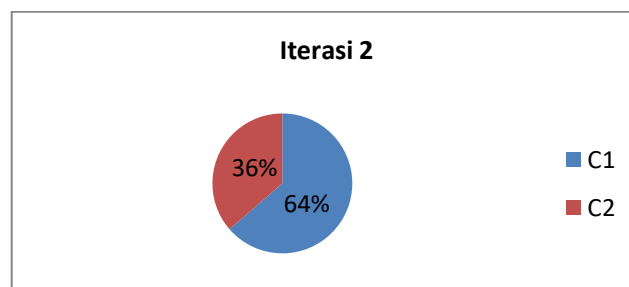
Klaster 1	Klaster 2
2,080192299	3,098935301
0,568594759	5,226327583
3,464852089	1,932097306
1,089082182	0

5,788307179	2,644636081
4,553965305	7,490026702
4,840847033	6,464487605
3,644447832	4,716015267
3,891040992	5,682508249
0,291204396	6,737692483
7,04482789	5,56493486
0	3,198280788
2,364846718	
4,324164659	
4,156657311	
1,193649865	
2,571867026	
1,400357097	
4,418653641	
8,036896167	
3,938857702	
69,66331214	52,75594223
122,4192544	

Tabel 6.Jumlah Distance Iterasi ke-2

Klaster 1	Klaster 2
2,357222942	5,014897806
0	6,762847034
2,908693865	0
1,62111073	1,932097306
5,224653099	1,64477962
4,488128786	8,191294159
4,318761397	8,333894648
3,332641595	3,795813483
3,97966079	7,593846193
0,85	5,058309994
6,512549424	6,978029808
0,568594759	3,694374101
2,580736329	
3,758257575	
3,81	
1,724093965	
2,716468295	
1,211651765	
4,036087214	
7,477031497	
4,050098764	
67,52644279	59,00018415
126,5266269	

Dari hasil pejumlahan data diatas distance baru – distance lama menghasilkan simpangan > 0, maka proses pengelompokan berhenti pada iterasi kedua dengan hasil simpangan 4,107372573



Gambar 2. Grafik hasil

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan maka dapat disimpulkan bahwa Algoritma K-Medoid pada pengelompokan perlakuan memilah sampah berdasarkan provinsi dapat diterapkan dengan baik. Data yang diolah menjadi 2 cluster yaitu cluster tingkat memilah sampah rendah (C1) dan cluster tingkat memilah sampah tinggi (C2). Dimana hasil penelitian ini menyimpulkan dari 33 provinsi di Indonesia bahwa cluster tingkat perilaku memilah sampah rendah (C1) diperoleh 22 provinsi yaitu Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kep. Bangka Belitung, Kep. Riau, DKI Jakarta, Jawa Timur, Banten, NTB, Kalimantan Barat, Kalimantan Selatan, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua, dan 12 provinsi lainnya termasuk dalam cluster tingkat memilah sampah tinggi (C2).

REFERENCES

- [1] B. Yulianto, "Partisipasi Pedagang Dalam Melakukan Pemilahan Sampah di Pasar Baru Kecamatan Tampan Kota Pekanbaru Participation Traders Separating Waste in Pasar Baru Tampan sub District Pekanbaru City," vol. 3, no. 18, pp. 69–72, 2016.
- [2] M. G. Sadewo, A. Eriza, A. P. Windarto, and D. Hartama, "Algoritma K-Means Dalam Mengelompokkan Desa / Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama Berdasarkan Provinsi," pp. 754–761, 2019.
- [3] A. P. Windarto, P. Studi, S. Informasi, and D. Mining, "Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering," vol. 16, no. 4, pp. 348–357, 2017.
- [4] A. P. Windarto, "Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering," *Techno.COM*, vol. 16, no. 4, pp. 348–357, 2017.
- [5] D. F. Pramesti, M. T. Furqon, and C. Dewi, "Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data Potensi Kebakaran Hutan / Lahan Berdasarkan Persebaran Titik Panas (Hotspot)," *J-ptiik*, vol. 1, no. 9, pp. 723–732, 2017.
- [6] W. A. Triyanto, "Algoritma K-Medoids Untuk Penentuan Strategi Pemasaran," vol. 6, no. 1, pp. 183–188, 2015.
- [7] S. Defiyanti, M. Jajuli, and N. Rohmawati, "Optimalisasi K-MEDOID dalam Pengklasteran Mahasiswa Pelamar Beasiswa dengan CUBIC CLUSTERING CRITERION," *J. Nas. Teknol. dan Sist. Inf.*, vol. 3, no. 1, pp. 211–218, 2017.
- [8] I. Parlina, A. P. Windarto, A. Wanto, and M. R. Lubis, "MEMANFAATKAN ALGORITMA K-MEANS DALAM MENENTUKAN PEGAWAI YANG LAYAK MENGIKUTI ASESSMENT CENTER," vol. 3, no. 1, pp. 87–93, 2018.
- [9] H. Zayuka, S. M. Nasution, and Y. Purwanto, "Perancangan Dan Analisis Clustering Data Menggunakan Metode K-Medoids Untuk Berita Berbahasa Inggris Design and Analysis of Data Clustering Using K-Medoids Method For English News," *e-Proceeding Eng.*, vol. 4, no. 2, pp. 2182–2190, 2017.