MENGKLUSTER JUMLAH KABUPATEN/KOTA YANG MELAKSANAKAN KAWASAN TANPA ROKOK (KTR) DI 50% SEKOLAH MENURUT PROVINSI DENGAN K-MEDOIDS

Erika Febrianti¹, Rahmat W Sembiring², Drs. Suhada³

¹Mahasiswa STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

^{2,3}Dosen STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Email: ¹Erikafebrianti@gmail.com, ²rahmatwsi@gmail.com, ³suhada.atb@gmail.com

Abstrak

Kawasan tanpa rokok (KTR) merupakan ruangan atau area yang dinyatakan dilarang untuk melakukan kegiatan merokok, kegiatan memproduksi, menjual, mengiklankan, promosi, dan atau mempromosikan produk tembakau. Masalah merokok sampai saat ini masih menjadi masalah nasional yang perlu secara terus menerus diupayakan penanggulangannya, karena menyangkut berbagai aspek permasalahan dalam kehidupan, yaitu aspek ekonomi, sosial, politik, utamanya aspek kesehatan. Diperkirakan lebih dari 40,3 juta anak tinggal bersama dengan perokok dan terpapar pada asap rokok di lingkungannya dan disebut sebagai perokok pasif. Sedangkan kita tahu bahwa anak yang terpapar asap rokok dapat mengalami peningkatan risiko terkena Bronkitis, Pneumonia, infeksi telinga tengah, Asma, serta kelambatan pertumbuhan paru-paru. Kerusakan kesehatan dini ini dapat menyebabkan kesehatan yang buruk pada masa dewasa. Orang dewasa bukan perokok pun yang terus-menerus terpapar juga akan mengalami peningkatan risiko Kanker Paru dan jenis kanker lainnya. Algoritma K-Medoids Clustering adalah salah satu algoritma yang digunakan untuk mengelompokkan data. Data yang digunakan yaitu data 34 provinsi yang melaksanakan kawasan tanpa rokok yang diperoleh dari dinas kesehatan RI. Data tersebut akan dikelompokkan menjadi dua cluster yaitu cluster tinggi dan rendah. Hasil yang diperoleh diharapkan dapat menjadi masukan kepada dinas terkait dalam menerapkan kawasan tanpa rokok.

Kata Kunci : Data Mining, K-Medoids, Clustering, kawasan tanpa rokok.

Abstract

A non-smoking area (KTR) is a room or area that is declared prohibited from engaging in smoking, producing, selling, advertising, promoting, or promoting tobacco products. The problem of smoking is still a national problem that needs to be continuously addressed, because it involves various aspects of problems in life, namely economic, social, political aspects, especially health aspects. It is estimated that more than 40.3 million children live with smokers and are exposed to cigarette smoke in their environment and are referred to as passive smokers. Meanwhile, we know that children who are exposed to secondhand smoke can have an increased risk of developing bronchitis, pneumonia, middle ear infections, asthma, and slowed lung growth. This early health damage can cause poor health in adulthood. Even non-smoker adults who are constantly exposed will also have an increased risk of lung cancer and other types of cancer. K-Medoids Clustering Algorithm is one algorithm used to group data. The data used are the data of 34 provinces implementing non-smoking areas obtained from the Indonesian Health Office. The data will be grouped into two clusters namely high and low clusters. The results obtained are expected to be input to the relevant agencies in implementing the non-smoking area.

Keywords : Data Mining, K-Medoids, Clustering, non smoking area.

1. PENDAHULUAN

Kawasan Tanpa Rokok (KTR) adalah ruangan atau area yang dinyatakan dilarang untuk kegiatan merokok atau kegiatan memproduksi, menjual, mengiklankan dan/atau mempromosikan produk tembakau. Oleh karena itu semua tempat yang telah ditetapkan sebagai KTR harus bebas dari asap rokok, penjualan, produksi, promosi dan sponsor rokok.

Pemerintah melalui UU No. 36 Tahun 2009 tentang Kesehatan dan PP No. 109 Tahun 2012 Tentang Pengamanan Bahan yang Mengandung Zat Adiktif Berupa Produk Tembakau Bagi Kesehatan telah mewajibkan pemerintah daerah untuk menetapkan KTR di wilayahnya masing-masing melalui Peraturan Daerah (Perda) atau peraturan perundang-undangan daerah lainnya. KTR ini meliputi: fasilitas pelayanan kesehatan, tempat proses belajar mengajar, tempat anak bermain, tempat ibadah, angkutan umum, tempat kerja, tempat umum dan tempat lain yang ditetapkan.

Oleh karena itu, diperlukan komitmen bersama dari lintas sektor dan berbagai elemen masyarakat ini akan sangat berpengaruh pada penerapan KTR. Di Indonesia kawasan tanpa rokok ini menjadi salah satu masalah yang besar dimata masyarakat khususnya bagi remaja yang masih sekolah. Oleh karena itu, perlu adanya suatu sistem untuk mengelompokkan provinsi-provinsi yang ada di Indonesia menjadi dua *cluster* yaitu *cluster* tinggi dan *cluster* rendah, dimana provinsi yang masuk ke dalam *cluster* tinggi menjadi pusat perhatian bagi pemerintah.

Data mining merupakan kegiatan menemukan pola yang menarik dari data dalam jumlah besar, data dapat disimpan dalam *database*, data *warehouse*, atau penyimpanan informasi lainnya. Data mining berkaitan dengan bidang ilmu – ilmu lain, seperti *database system*, data *warehousing*, statistik, *machine learning*, information retrieval, dan komputasi tingkat tinggi. Selain itu, data mining didukung oleh ilmu lain seperti *neural network*, pengenalan pola, spatial data analysis, image database, signal processing (Defiyanti sofi, dkk., 2017).

Algoritma PAM sering dikenal dengan algoritma K-Medoids merupakan sebuah algoritma yang merepresentasikan *cluster* yang dibentuk menggunakan *medoids*. Proses pembentukan *cluster* dimulai dengan

menentukan k objek dari *dataset* secara acak sebagai *medoid*, selanjutnya hitung *cost* setiap objek *non-medoid* dengan k objek, *cost* terkecil setiap objek *non-medoid* terhadap *medoid* akan masuk dalam *cluster* dimana *medoid* tersebut berada (Chrisnanto and Abdillah, 2015. kontribusi pemerintah provinsi untuk meningkatkan fasilitas dikawasan sekolah yang belum memiliki kawasan tanpa rokok (KTR).

2. TEORITIS

2.1. Data Mining

Data Mining merupakan proses mempekerjakan satu atau lebih teknik pembelajaran *computer (machine learning)* untuk menganalisis dan mengekstraksi pengetahuan (knowledge) secara otomatis [1]. Dalam pengertian lain mengatakan bahwa Data Mining merupakan proses ataupun kegiatan untuk mengumpulkan data yang berukuran besar kemudian mengekstraksi data tersebut menjadi informasi – informasi yang nantinya dapat digunakan [2].

2.2. Algoritma K-Medoids

Menurut [3] *Algoritma Partition Around Medoids (PAM)* atau dikenal juga dengan *K-medoids* dikembangkan oleh Leonard Kaufman dan Peter J. Rousseeuw pada tahun 1987. *Algoritma PAM* termasuk metode *partitioning clustering* untuk mengelompokkan sekumpulan n objek menjadi sekumpulan k *cluster*. Representasi *cluster* pada PAM adalah objek dari sekumpulan objek yang mewakili *cluster*, disebut *medoid*. *Cluster* dibangun dengan menghitung kedekatan yang dimiliki antara *medoid* dengan objek *non-medoid* menggunakan *distance measure*. Jadi metode partisi masih dapat dilakukan berdasarkan prinsip meminimalkan jumlah dari ketidakmiripan antara setiap objek dan *medoid* yang sesuai.

Menurut (Kaur, dkk., 2014) dikutip oleh [4] *K-Medoids* atau *Partitioning Around Medoids (PAM)* adalah *algoritma clustering* yang mirip dengan *K-Means*. Perbedaan dari kedua *algoritma* ini yaitu *algoritma K-Medoids* atau PAM menggunakan objek sebagai perwakilan (*medoid*) sebagai pusat *cluster* untuk setiap *cluster*, sedangkan *K-Means* menggunakan nilai rata-rata (*mean*) sebagai pusat *cluster*.

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 dari distribusi data. Kelebihan lainnya yaitu hasil proses *clustering* tidak bergantung pada urutan masuk *dataset* [4].

K-Medoids atau *Partitioning Around Medoids (PAM)* adalah *algoritma clustering* yang hampir sama dengan *algoritma K-Means*. Yang membedakan kedua *algoritma* ini adalah *K-Medoids* atau PAM menggunakan objek sebagai perwakilan (*medoid*) sebagai pusat *cluster* untuk setiap *cluster*, sedangkan *K-Means* menggunakan nilai rata-rata (*mean*) sebagai pusat *cluster*. Kelebihan dari *K-Medoids* yaitu mengatasi kelemahan pada *algoritma K-Means* yang rentan terhadap *noise* dan *outlier*, dimana objek dengan nilai yang besar yang memungkinkan menyimpang pada dari distribusi data. Kelebihan lainnya adalah hasil *clustering* tidak bergantung pada urutan masuk data set [4]. Langkah-langkah *algoritma K-Medoids* adalah (Zayuka, Nasution, & Purwanto, 2017):

1. Pilih poin k sebagai inisial *centroid* / nilai tengah (*medoids*) sebanyak k *cluster*.
2. Cari semua poin yang terdekat dengan *medoid*, dengan cara menghitung jarak vektor antar dokumen dengan menggunakan *EuclidianDistance*. Rumusnya adalah sebagai berikut:

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

Dimana:

$d(x,y)$ = jarak antara data ke-i dan data ke-j

x_{i1} = nilai atribut ke satu dari data ke-i

y_{j1} = nilai atribut ke satu dari data ke-j

n = jumlah atribut yang digunakan

3. Secara acak, pilih poin yang bukan *medoid*.
4. Hitung total jarak antar *medoid*.
5. Jika TD baru < TD awal, tukar posisi *medoids* dengan *medoid* baru, jadilah *medoid* yang baru.
6. Ulangi langkah 2-5 sampai *medoid* tidak berubah.

2.4. Teknik Pengumpulan Data

Penelitian ini dilakukan berdasarkan provinsi di Indonesia dengan mengambil data langsung dari situs Departemen kesehatan dengan alamat url <https://www.depkes.go.id/>. Sampel data yang digunakan dalam penelitian ini adalah nama-nama provinsi yang ada di Indonesia yang terdiri dari 34 provinsi. Data yang diambil dari Departemen Kesehatan.

Ada dua metode dalam teknik pengumpulan data yaitu:

1. Metode Sensus, yaitu teknik penentuan sampel bila semua anggota populasi digunakan sebagai sampel. Hal ini sering dilakukan bila jumlah populasinya relatif kecil, kurang dari 30 orang, dimana semua anggota populasi dijadikan sampel.

2. Metode Sampling, yaitu teknik yang dilakukan untuk menentukan sampel. Jadi, sebuah penelitian yang baik haruslah memperhatikan dan menggunakan sebuah teknik dalam menetapkan sampel yang akan diambil sebagai subjek penelitian.

Dalam memperoleh data penulis melakukan teknik sampling yaitu mengambil sebagai data untuk diteliti. Adapun cara mendapatkan datanya adalah :

1. Metode Observasi, melihat serta mempelajari permasalahan yang ada dilapangan yang berkaitan dengan objek yang diteliti, yaitu informasi mengenai kabupaten/kota yang melaksanakan kawasan tanpa rokok di sekolah menurut provinsi
2. Wawancara
Proses memperoleh keterangan untuk tujuan penelitian dengan cara tanya jawab sambil bertatap muka antara pewawancara dengan responden atau orang yang diwawancarai.
3. Studi kepustakaan (*Library Research*) yaitu memanfaatkan perpustakaan, buku, prosiding atau jurnal sebagai media untuk bahan referensi dalam menentukan parameter yang digunakan dalam penelitian.

3. ANALISA DAN PEMBAHASAN

3.1. Pengolahan Data

Beberapa tahapan dalam proses perhitungan manual menggunakan algoritma *k-medoids* pada kawasan tanpa rokok di Indonesia sebagai berikut :

Tabel 1. Data KawasanTanpaRokok

No.	Provinsi	Jumlah Kabupaten/Kota	Implementasi KTR Pada 50% Sekolah	Persentase
1.	Aceh	23	4	17.391
2.	Sumatera Utara	33	4	12.121
3.	Sumatera Barat	19	11	57.895
4.	Riau	12	3	25
5.	Jambi	11	4	36.364
6.	Sumatera selatan	17	7	41.176
7.	Bengkulu	10	2	20
8.	Lampung	15	6	40
9.	Kep.Bangka Belitung	7	3	42.857
10.	Kepulauan Riau	7	3	42.857
11.	DKI Jakarta	6	6	100
12.	Jawa Barat	27	9	33.333
13.	Jawa Tengah	35	6	17.143
14.	Yogyakarta	5	5	100
15.	JawaTimur	38	6	15.789
16.	Banten	8	4	50
17.	Bali	9	9	100
18.	NTB	10	5	50
19.	NTT	22	2	9.091
20.	Kalimantan Barat	14	5	35.714
21.	Kalimantan Tengah	14	3	21.429
22.	Kalimantan Selatan	13	5	38.462
23.	Kalimantan Timur	10	6	60
24.	Kalimantan Utara	5	2	40
25.	Sulawesi Utara	15	5	33.333
26.	Sulawesi Tengah	13	2	15.385
27.	Sulawesi Selatan	24	7	29.167
28.	Sulawesi Tenggara	17	2	11.765
29.	Gorontalo	6	3	50
30.	Sulawesi Barat	6	3	50
31.	Maluku	11	4	36.364
32.	Maluku Utara	10	3	30
33.	Papua Barat	13	2	15.385
34.	Papua	29	3	10.345

Menentukan jumlah *cluster* (k) dari n objek adalah $2cluster$.

1. Mengasumsikan *centroid* awal yang telah ditentukan seperti pada tabel 1.

Tabel 2. *Medoids*

Nama	Provinsi	Jumlah Kabupaten/Kota	Implementasi KTR Pada 50% Sekolah	%
C1	Jawa Barat	27	9	33,333
C2	Jambi	11	4	36,364

Menempatkan objek-objek *non medoids* kedalam *cluster* yang paling dekat dengan *medoids* berdasarkan jarak *Euclidean*. Berikut merupakan perhitungan jarak pada data kawasan tanpa rokok:

$$d_{Aceh,C1} = \sqrt{(23-27)^2 + (4-9)^2 + (17,391-33,333)^2} = 17,1798$$

$$d_{Aceh,C2} = \sqrt{(23-11)^2 + (4-4)^2 + (17,391-36,364)^2} = 22,4493$$

$$d_{Sumut,C1} = \sqrt{(33-27)^2 + (4-9)^2 + (17,391-33,333)^2} = 22,6041$$

$$d_{Sumut,C2} = \sqrt{(23-11)^2 + (4-4)^2 + (17,391-36,364)^2} = 32,7371$$

$$d_{Sumbar,C1} = \sqrt{(19-27)^2 + (11-9)^2 + (57,895-33,333)^2} = 25,9093$$

$$d_{Sumbar,C2} = \sqrt{(19-11)^2 + (11-4)^2 + (57,895-36,364)^2} = 24,0121$$

$$d_{Riau,C1} = \sqrt{(12-27)^2 + (3-9)^2 + (25-33,333)^2} = 18,1779$$

$$d_{Riau,C2} = \sqrt{(12-11)^2 + (3-4)^2 + (25-36,364)^2} = 11,4516$$

$$d_{Jambi,C1} = \sqrt{(11-27)^2 + (4-9)^2 + (36,364-33,333)^2} = 17,0348$$

$$d_{Jambi,C2} = \sqrt{(11-11)^2 + (4-4)^2 + (36,364-36,364)^2} = 0$$

Untuk melihat hasil perhitungan jarak ke setiap *medoids*, dapat dilihat pada tabel 3 berikut:

Tabel 3. Hasil Perhitungan Jarak ke Setiap *Medoids*

No.	Data ke i	Jarak ke medoids		Jarak Terdekat	Cluster yang diikuti
		C1	C2		
1.	Aceh	17,17985	22,44938	17,17985	1
2.	Sumatera Utara	22,60418	32,73718	22,60418	1
3.	Sumatera Barat	25,9093	24,01216	24,01216	2
4.	Riau	18,17798	11,45166	11,45166	2
5.	Jambi	17,03487	0	0	2
6.	Sumatera selatan	12,86517	8,255625	8,255625	2
7.	Bengkulu	22,71055	16,51607	16,51607	2
8.	Lampung	14,05165	5,763722	5,763722	2
9.	Kep. Bangka Belitung	22,95009	7,691492	7,691492	2
10.	Kepulauan Riau	22,95009	7,691492	7,691492	2
11.	DKI Jakarta	69,96062	63,86345	63,86345	2
12.	Jawa Barat	0	17,03487	0	1
13.	Jawa Tengah	18,30618	30,8131	18,30618	1
14.	Yogyakarta	70,31706	63,92605	63,92605	2
15.	Jawa Timur	20,92348	34,00486	20,92348	1
16.	Banten	25,7641	13,96211	13,96211	2
17.	Bali	69,05425	63,86345	63,86345	2
18.	Nusa Tenggara Barat	24,14102	13,70914	13,70914	2
19.	Nusa Tenggara Timur	25,72304	29,47569	25,72304	1
20.	Kalimantan Barat	13,8083	3,22839	3,22839	2
21.	kalimantan Tengah	18,62002	15,26611	15,26611	2
22.	Kalimantan Selatan	15,43718	3,066204	3,066204	2
23.	Kalimantan Timur	31,76679	23,74154	23,74154	2
24.	Kalimantan Utara	24,03017	7,295238	7,295238	2
25.	Sulawesi Utara	12,64911	5,11732	5,11732	2
26.	Sulawesi Tengah	23,81451	21,16881	21,16881	2
27.	Sulawesi Selatan	5,509588	15,15905	5,509588	1
28.	Sulawesi Tenggara	24,78263	25,39903	24,78263	1
29.	Gorontalo	27,47342	14,55818	14,55818	2
30.	Sulawesi Barat	27,47342	14,55818	14,55818	2
31.	Maluku	17,03487	0	0	2
32.	Maluku Utara	18,33327	6,51924	6,51924	2
33.	Papua Barat	23,81451	21,16881	21,16881	2
34.	Papua	23,84215	31,6542	23,84215	1
Jumlah		829,0134	675,1218		
Total Cost		1504,135			

Mengansumsikan objek yang telah ditentukan sebagai *non medoids* sebagai berikut :

Tabel 4. *Non Medoids*

Nama	Provinsi	Jumlah Kabupaten/Kota	Implementasi KTR Pada 50% Sekolah	%
C1	Sulawesi Selatan	24	7	29,167
C2	Jawa Timur	38	6	15,789

Menempatkan kembali objek-objek *non medoids* kedalam *cluster* yang paling dekat dengan *medoids* berdasarkan jarak *Euclidean*. Berikut adalah perhitungan jarak pada data kawasan tanpa rokok:

$$d_{Aceh,C1} = \sqrt{(23-24)^2 + (4-7)^2 + (17,391-29,167)^2} = 15,2173$$

$$d_{Aceh,C2} = \sqrt{(23-38)^2 + (4-6)^2 + (17,391-15,789)^2} = 12,1932$$

$$d_{Sumut,C1} = \sqrt{(33-24)^2 + (4-7)^2 + (17,391-29,167)^2} = 19,5081$$

$$d_{Sumut,C2} = \sqrt{(23-38)^2 + (4-6)^2 + (17,391-15,789)^2} = 6,5156$$

$$d_{Sumbar,C1} = \sqrt{(19-24)^2 + (11-7)^2 + (57,895-29,167)^2} = 29,4329$$

$$d_{Sumbar,C2} = \sqrt{(19-38)^2 + (11-6)^2 + (57,895-15,789)^2} = 46,4641$$

$$d_{Riau,C1} = \sqrt{(12-24)^2 + (3-7)^2 + (25-29,167)^2} = 13,3178$$

$$d_{Riau,C2} = \sqrt{(12-38)^2 + (3-6)^2 + (25-15,789)^2} = 27,7460$$

$$d_{Jambi,C1} = \sqrt{(11-24)^2 + (4-7)^2 + (36,364-29,167)^2} = 15,1590$$

$$d_{Jambi,C2} = \sqrt{(11-38)^2 + (4-6)^2 + (36,364-15,789)^2} = 34,0048$$

Untuk melihat hasil perhitungan jarak ke setiap *non medoids*, dapat dilihat pada tabel 5 sebagai berikut:

Tabel 5. Hasil Perhitungan Jarak ke Setiap *Non Medoids*

No.	Data Ke I	Jarak Ke Medoids		Jarak terdekat	Cluster Yang Diikuti
		C1	C2		
1.	Aceh	15,21731	12,1932	12,1932	2
2.	Sumatera Utara	19,5081	6,515691	6,515691	2
3.	Sumatera Barat	29,43294	46,46413	29,43294	1
4.	Riau	13,3178	27,74604	13,3178	1
5.	Jambi	15,15905	34,00486	15,15905	1
6.	Sumatera selatan	13,90022	32,9621	13,90022	1
7.	Bengkulu	17,46522	28,59602	17,46522	1
8.	Lampung	14,11927	33,3942	14,11927	1
9.	Kep. Bangka Belitung	22,19045	41,2635	22,19045	1
10.	Kepulauan Riau	22,19045	41,2635	22,19045	1
11.	DKI Jakarta	73,09113	90,08603	73,09113	1
12.	Jawa Barat	5,509588	20,92348	5,509588	1
13.	Jawa Tengah	16,32717	3,2914	3,2914	2
14.	Yogyakarta	73,36425	90,4516	73,36425	1
15.	Jawa Timur	19,38997	0	0	2
16.	Banten	26,43887	45,5455	26,43887	1
17.	Bali	72,43144	89,11505	72,43144	1
18.	Nusa Tenggara Barat	25,17963	44,21982	25,17963	1
19.	Nusa Tenggara Timur	20,78571	17,80065	17,80065	2
20.	Kalimantan Barat	12,11871	31,20906	12,11871	1
21.	kalimantan Tengah	13,26185	24,83565	13,26185	1
22.	Kalimantan Selatan	14,5395	33,76485	14,5395	1
23.	Kalimantan Timur	33,87734	52,33175	33,87734	1
24.	Kalimantan Utara	22,43555	41,12387	22,43555	1

25.	Sulawesi Utara	10,11709	28,94464	10,11709	1
26.	Sulawesi Tengah	18,32876	25,3212	18,32876	1
27.	Sulawesi Selatan	0	19,38997	0	1
28.	Sulawesi Tenggara	19,4121	21,75299	19,4121	1
29.	Gorontalo	27,82111	46,94031	27,82111	1
30.	Sulawesi Barat	27,82111	46,94031	27,82111	1
31.	Maluku	15,15905	34,00486	15,15905	1
32.	Maluku Utara	14,58403	31,54287	14,58403	1
33.	Papua Barat	18,32876	25,3212	18,32876	1
34.	Papua	19,88134	10,93788	10,93788	2
Jumlah		779,6808	1183,222		
Total Cost		1962,903			

2. Menghitung nilai S dengan persamaan (1). sebagai berikut:

S = total *cost* baru – total *cost* lama

Dengan : S = Selisih

Total *cost* baru = total *cost* non *medoids*

Total *cost* lama = total *cost* untuk *medoids*

Sehingga diperoleh:

$S = 1962,903 - 1504,135 = 458,767$

3. Karena nilai $S > 0$ maka proses pengklasteran dihentikan. Sehingga diperoleh anggota tiap kluster sebagai berikut.

Tabel 6. Hasil Pengklasteran Data dengan *K-Medoids Clustering*

No.	Data Ke I	Jarak Ke Medoids		Jarak Terdekat	Cluster Yang Diikuti
		C1	C2		
1.	Aceh	15.21731	12.1932	12.1932	2
2.	Sumatera Utara	19.5081	6.515691	6.515691	2
3.	Sumatera Barat	29.43294	46.46413	29.43294	1
4.	Riau	13.3178	27.74604	13.3178	1
5.	Jambi	15.15905	34.00486	15.15905	1
6.	Sumatera selatan	13.90022	32.9621	13.90022	1
7.	Bengkulu	17.46522	28.59602	17.46522	1
8.	Lampung	14.11927	33.3942	14.11927	1
9.	Kep. Bangka Belitung	22.19045	41.2635	22.19045	1
10.	Kepulauan Riau	22.19045	41.2635	22.19045	1
11.	DKI Jakarta	73.09113	90.08603	73.09113	1
12.	Jawa Barat	5.509588	20.92348	5.509588	1
13.	Jawa Tengah	16.32717	3.2914	3.2914	2
14.	Yogyakarta	73.36425	90.4516	73.36425	1
15.	Jawa Timur	19.38997	0	0	2
16.	Banten	26.43887	45.5455	26.43887	1
17.	Bali	72.43144	89.11505	72.43144	1
18.	Nusa Tenggara Barat	25.17963	44.21982	25.17963	1
19.	Nusa Tenggara Timur	20.78571	17.80065	17.80065	2
20.	Kalimantan Barat	12.11871	31.20906	12.11871	1
21.	kalimantan Tengah	13.26185	24.83565	13.26185	1
22.	Kalimantan Selatan	14.5395	33.76485	14.5395	1
23.	Kalimantan Timur	33.87734	52.33175	33.87734	1
24.	Kalimantan Utara	22.43555	41.12387	22.43555	1
25.	Sulawesi Utara	10.11709	28.94464	10.11709	1
26.	Sulawesi Tengah	18.32876	25.3212	18.32876	1
27.	Sulawesi Selatan	0	19.38997	0	1
28.	Sulawesi Tenggara	19.4121	21.75299	19.4121	1
29.	Gorontalo	27.82111	46.94031	27.82111	1
30.	Sulawesi Barat	27.82111	46.94031	27.82111	1
31.	Maluku	15.15905	34.00486	15.15905	1

32.	Maluku Utara	14.58403	31.54287	14.58403	1
33.	Papua Barat	18.32876	25.3212	18.32876	1
34.	Papua	19.88134	10.93788	10.93788	2

3.2. TAMPILAN RAPIDMINER

Berdasarkan gambar 1 hasil *clustering* maka didapat grafik hasil dari rapidminer sebagai berikut :

Gambar 1. Grafik *Clustering*

Berdasarkan gambar 3.10. dapat diketahui bahwa hasil perhitungan dengan menggunakan *tools rapidminer* didapat jumlah provinsi di cluster_0(C1 pada perhitungan manual) sebanyak 6 provinsi dan 28 provinsi di cluster_1(C2 pada perhitungan manual). Berikut merupakan tampilan *cluster model* berupa teks yang ada dalam aplikasi *RapidMiner 5.3*. dimana tampilan ini adalah penjelasan dari *cluster* yang terbentuk. Dapat dilihat pada gambar 2:

Gambar 2. *Cluster Model* Kawasan Tanpa Rokok

Dalam melakukan *validasi* data perhitungan algoritma harus menghasilkan hasil akhir berupa pengelompokkan dengan *2cluster*, serta data yang digunakan merupakan data yang *valid* dan sama dengan yang dipakai pada *tools Rapidminer 5.3*. Berikut ditampilkan hasil yang didapatkan dari perhitungan algoritma dan pengujian pada *Rapidminer 5.3*, pada table 7 dibawah ini.

Tabel 7. Perbandingan Hasil Akhir *K-Medoids*

Provinsi	Perhitungan Algoritma <i>K-Medoids</i>	Pengujian Dengan <i>Rapidminer 5.3</i>
Aceh	cluster_1	cluster_1
Sumatera utara	cluster_1	cluster_1
Sumatera barat	cluster_2	cluster_2
Riau	cluster_2	cluster_2
Jambi	cluster_2	cluster_2
Sumatera selatan	cluster_2	cluster_2
Bengkulu	cluster_2	cluster_2
Lampung	cluster_2	cluster_2
Kep. Bangka Belitung	cluster_2	cluster_2
Kep. Riau	cluster_2	cluster_2
DKI Jakarta	cluster_2	cluster_2
Jawa barat	cluster_2	cluster_2
Jawa tengah	cluster_1	cluster_1
DIYogyakarta	cluster_2	cluster_2

Provinsi	Perhitungan Algoritma <i>K-Medoids</i>	Pengujian Dengan <i>Rapidminer 5.3</i>
Jawa timur	cluster_1	cluster_1
Banten	cluster_2	cluster_2
Bali	cluster_2	cluster_2
Nusa tenggara barat	cluster_2	cluster_2
Nusa tenggara timur	cluster_1	cluster_1
Kalimantan barat	cluster_2	cluster_2
Kalimantan tengah	cluster_2	cluster_2
Kalimantan selatan	cluster_2	cluster_2
Kalimantan timur	cluster_2	cluster_2
Kalimantan utara	cluster_2	cluster_2
Sulawesi utara	cluster_2	cluster_2
Sulawesi tengah	cluster_2	cluster_2
Sulawesi selatan	cluster_2	cluster_2
Sulawesi tenggara	cluster_2	cluster_2
Gorontalo	cluster_2	cluster_2
Sulawesi barat	cluster_2	cluster_2
Maluku	cluster_2	cluster_2
Maluku utara	cluster_2	cluster_2
Papua barat	cluster_2	cluster_2
Papua	cluster_1	cluster_1
Hasil	cluster_1 cluster_2	5 29

4. KESIMPULAN

Adapun beberapa kesimpulan dari penelitian ini adalah :

1. Kawasan Tanpa Rokok (KTR) merupakan ruangan atau area yang dinyatakan dilarang untuk kegiatan merokok atau kegiatan memproduksi, menjual, mengiklankan dan / atau mempromosikan produk tembakau. Oleh karena itu semua tempat yang telah ditetapkan sebagai Kawasan Tanpa Rokok (KTR) harus bebas dari asap rokok, penjualan, produksi, promosi dan sponsor rokok.
2. *K-Medoids* atau *Partitioning Around Medoids* (PAM) merupakan algoritma *clustering* yang hampir sama dengan algoritma *K-Means*. Yang membedakan kedua algoritma ini adalah *K-Medoids* atau PAM menggunakan objek sebagai perwakilan (*medoid*) sebagai pusat *cluster* untuk setiap *cluster*, sedangkan *K-Means* menggunakan nilai rata-rata (*mean*) sebagai pusat *cluster*. Kelebihan dari *K-Medoids* yaitu mengatasi kelemahan pada algoritma *K-Means* yang rentan terhadap *noise* dan *outlier*, dimana objek dengan nilai yang besar yang memungkinkan menyimpang pada dari distribusi data.
3. Data yang digunakan diperoleh dari Depkes RI yang dapat diakses di www.depkes.go.id . Setelah melakukan cleaning data, data tersebut diolah dengan menggunakan algoritma *K-Medoids* dan dilakukan pengujian data menggunakan software rapidminer. hasil yang diperoleh dari perhitungan manual dan rapidminer menunjukkan hasil yang sama.

5. REFERENCES

- [1] M. G. Sadewo, A. P. Windarto, and D. Hartama, "Penerapan Datamining Pada Populasi Daging Ayam Ras Pedaging Di Indonesia Berdasarkan Provinsi Menggunakan K-Means," *InfoTekJar (Jurnal Nas. Inform. dan Teknol. Jaringan)*, vol. 2, no. 1, pp. 60–67, 2017.
- [2] A. Saleh, "PENERAPAN DATA MINING DENGAN METODE KLASIFIKASI NAÏVE BAYES UNTUK MEMPREDIKSI KELULUSAN MAHASISWA DALAM MENGIKUTI ENGLISH PROFICIENCY TEST (Studi Kasus : Universitas Potensi Utama) PENERAPAN DATA MINING DENGAN METODE KLASIFIKASI NAÏVE MENGIKUTI ENGLISH ," no. February 2015, 2016.
- [3] B. D. Satoto, B. K. Khotimah, and Iswati, "PENDIDIKAN MENGGUNAKAN PARTITION AROUND MEDOIDS DAN," vol. 1, no. 1, pp. 17–24, 2015.
- [4] 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. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 9, pp. 723–732, 2017.
- [5] H. Zayuka, S. M. Nasution, and Y. Purwanto, "Perancangan Dan Analisis Clustering Data Menggunakan Metode K-Medoids Untuk Berita Berbahasa Inggris," *e-Proceeding Eng. Vol.4, No.2 Agustus 2017*, vol. 4, no. 2, pp. 1–9, 2017.