
PENERAPAN ALGORITMA CLUSTERING DALAM MENGELOMPOKKAN BANYAKNYA DESA/KELURAHAN MENURUT UPAYA ANTISIPASI/ MITIGASI BENCANA ALAM MENURUT PROVINSI DENGAN K-MEANS

Mhd Gading Sadewo, Agus Perdana Windarto, Anjar Wanto

STIKOM Tunas Bangsa Pematangsiantar, Indonesia
Email: Gading.sadewooo@gmail.com

Abstrak

Bencana alam adalah suatu peristiwa alam yang mengakibatkan dampak besar bagi populasi manusia. Berlokasi di Cincin Api Pasifik (wilayah dengan banyak aktivitas tektonik), Indonesia harus terus menghadapi resiko letusan gunung berapi, gempa bumi, banjir, tsunami. Penerapan Algoritma Clustering Dalam Mengelompokkan Banyaknya Desa/Kelurahan Menurut Upaya Antisipasi/Mitigasi Bencana Alam Menurut Provinsi Dengan K-Means. Sumber data penelitian ini dikumpulkan berdasarkan dokumen-dokumen keterangan Banyaknya Desa/Kelurahan Menurut Upaya Antisipasi/Mitigasi Bencana Alam yang dihasilkan oleh Badan Pusat Statistik Nasional. Data yang digunakan dalam penelitian ini adalah data provinsi yang terdiri dari 34 provinsi. Variable yang digunakan ada 4 yaitu Sistem Peringatan Dini Bencana Alam, Sistem Peringatan Dini Tsunami, Perlengkapan Keselamatan, Jalur Evakuasi. Data akan diolah dengan melakukan clustering dalam 3 clushter yaitu clushter tingkat antisipasi/mitigasi tinggi, clushter tingkat antisipasi/mitigasi sedang dan tingkat antisipasi/mitigasi rendah. Hasil yang diperoleh dari proses penilaian berdasarkan indeks Desa/Kelurahan Menurut Upaya Antisipasi/Mitigasi Bencana Alam dengan 3 provinsi tingkat antisipasi/mitigasi tinggi yaitu Jawa Barat, Jawa Tengah, Jawa Timur, 9 provinsi tingkat antisipasi/mitigasi sedang, dan 22 provinsi lainnya termasuk tingkat antisipasi/mitigasi rendah. Hal ini dapat menjadi masukan kepada pemerintah, provinsi yang menjadi perhatian lebih pada Desa/Kelurahan Menurut Upaya Antisipasi/Mitigasi Bencana Alam Kesehatan berdasarkan cluster yang telah dilakukan.

Kata Kunci : *Data Mining, Bencana Alam, Clustering, K-Means*

Abstract

Natural disasters are natural events that have a large impact on the human population. Located on the Pacific Ring of Fire (an area with many tectonic activities), Indonesia must continue to face the risk of volcanic eruptions, earthquakes, floods, tsunamis. Application of Clustering Algorithm in Grouping the Number of Villages / Villages According to Anticipatory / Natural Disaster Mitigation Efforts by Province With K-Means. The source of this research data is collected based on documents that contain the number of villages / kelurahan according to natural disaster mitigation / mitigation efforts produced by the National Statistics Agency. The data used in this study is provincial data consisting of 34 provinces. There are 4 variables used, namely the Natural Disaster Early Warning System, Tsunami Early Warning System, Safety Equipment, Evacuation Line. The data will be processed by clustering in 3 clushter, namely clushter high level of anticipation / mitigation, clusters of moderate anticipation / mitigation levels and low anticipation / mitigation levels. The results obtained from the assessment process are based on the Village / Kelurahan index according to the Natural Disaster Anticipation / Mitigation Efforts with 3 provinces of high anticipation / mitigation levels, namely West Java, Central Java, East Java, 9 provinces of moderate anticipation / mitigation, and 22 other provinces including low anticipation / mitigation. This can be an input to the government, the provinces that are of greater concern to the Village / Village According to the Natural Health Disaster Mitigation / Mitigation Efforts based on the cluster that has been carried out.

Keywords: *Data Mining, Natural Disaster, Clustering, K-Means*

1. PENDAHULUAN

Bencana alam adalah suatu peristiwa alam yang mengakibatkan dampak besar bagi populasi manusia. Indonesia merupakan Negara yang sangat rawan dengan bencana alam seperti gempa bumi, tsunami, letusan gunung berapi, tanah longsor, banjir, dan puing beliuang. Sekitar 13 persen gunung berapi dunia yang berada di kepulauan Indonesia berpotensi menimbulkan bencana dengan intensitas dan kekuatan yang berbeda-beda. Penanggulangan bencana alam atau mitigasi adlah upaya berkelanjutan untuk mengurangi dampak bencana terhadap manusia dan harta benda. Persiapan menghadapi bencana alam termasuk semua kativitas yang dilakukan sebelum terdeteksinya tanda-tanda bencana agar bias memfasilitasi pemakaian sumber daya alam yang tersedia, meminta bantuan dan serta rencan rehabilitasi dalam cara dan kemungkinan yang paling baik. Kesiapan menghadapi bencana alam dimulai dari level komunikasi local. Jika sumber daya lokal kurang mencukupi, maka daerah tersebut dapat meminta bantuan ke tingkat nasional dan internasional. Banyak cabang ilmu komputer yang

dapat menyelesaikan permasalahan yang bersifat kompleks. Hal ini dibuktikan oleh beberapa penelitian dibidang datamining [1][2]–[5], bidang jaringan saraf tiruan [6], [7][6]–[9], dalam bidang sistem pendukung keputusan [10]–[14]. Berdasarkan penjelasan tersebut, peneliti menggunakan datamining dapat menyelesaikan masalah diatas.

Analisis klaster merupakan suatu teknik multivariate dengan tujuan utama mengelompokkan objek-objek berdasarkan karakteristik yang dimiliki. Sekarang ini analisis klaster telah banyak diaplikasikan dalam berbagai bidang yang ditulis dalam berbagai penelitian dan jurnal [1]. Dalam algoritma clustering konsep utama yang ditekankan adalah pencarian pusat cluster secara iteratif, dimana pusat cluster ditentukan berdasarkan jarak minimum setiap data pada pusat cluster [1]. Data yang digunakan dalam penelitian ini berdasarkan dokumen-dokumen yang dihasilkan oleh Badan Pusat Statistik Nasional melalui situs <https://www.bps.go.id>. Dalam hal ini peneliti mengangkat topik mengelompokkan desa/kelurahan menurut upaya antisipasi/mitigasi bencana alam menurut provinsi dimana proses algoritma yang dilakukan adalah clustering. Hasil dari cluster dapat dijadikan masukan bagi pemerintah agar provinsi yang masuk kedalam clushter rendah mendapat perhatian lebih dalam peningkatan sarana dan prasarana upaya antisipasi bencana. Proses clushter dibagi kedalam 3 (tiga) cluster yakni sarana kesehatan tertinggi, sarana kesehatan sedang dan sarana kesehatan rendah.

2. TEORITIS

2.1. Data Mining

Data mining yang juga dikenal dengan istilah *pattern recognition* merupakan suatu algoritma yang digunakan untuk pengolahan data guna menemukan pola yang tersembunyi dari data yang diolah. Data yang diolah dengan teknik data mining ini kemudian menghasilkan suatu pengetahuan baru yang bersumber dari data lama, hasil dari pengolahan data tersebut dapat digunakan dalam menentukan keputusan di masa depan[2][3].

2.2. Clustering

Analisis Pengelompokan / Clustering merupakan proses membagi data dalam suatu himpunan ke dalam beberapa kelompok yang kesamaan datanya dalam suatu kelompok lebih besar daripada kesamaan data tersebut dengan data dalam kelompok lain. Potensi clustering adalah dapat digunakan untuk mengetahui struktur dalam data yang dapat dipakai lebih lanjut dalam berbagai aplikasi secara luas seperti klasifikasi, pengolahan gambar, dan pengenalan pola[4].

2.3. K-Means

K-Means merupakan suatu algoritma yang digunakan dalam pengelompokkan secara pertisi yang memisahkan data ke dalam kelompok yang berbeda – berbeda. Algoritma ini mampu meminimalkan jarak antara data ke *cluster*nya. Pada dasarnya penggunaan algoritma ini dalam proses *clustering* tergantung pada data yang didapatkan dan konklusi yang ingin dicapai di akhir proses[5]. Sehingga dalam penggunaan algoritma kmeans terdapat aturan sebagai berikut [5] :

- a. Berapa jumlah *cluster* yang perlu dimasukkan.
- b. Hanya memiliki atribut bertipe numeric.

Pada dasarnya algoritma k-means hanya mengambil sebagian dari banyaknya komponen yang didapatkan untuk kemudian dijadikan pusat *cluster* awal, pada penentuan pusat *cluster* ini dipilih secara acak dari populasi data. Kemudian algoritma k-means akan menguji masing – masing dari setiap komponen dalam populasi data tersebut dan menandai komponen tersebut ke dalam salah satu pusat *cluster* yang telah didefinisikan sebelumnya tergantung dari jarak minimum antar komponen dengan tiap – tiap pusat *cluster*. Selanjutnya posisi pusat *cluster* akan dihitung kembali sampai semua komponen data digolongkan ke dalam tiap – tiap *cluster* dan terakhir akan terbentuk *cluster* baru[5].

3. ANALISA DAN PEMBAHASAN

3.1. Tahap Pengumpulan Data

Dalam Pemanfaatan Algoritma Clustering Dalam Mengelompokkan Banyaknya Desa/Kelurahan Menurut Upaya Antisipasi / Mitigasi Bencana Alam Menurut Provinsi, diperlukan data terkait tentang hal itu. Sumber data penelitian diperoleh dari data yang dikumpulkan berdasarkan dokumen-dokumen keterangan dihasilkan oleh Direktorat Jenderal melalui situs <https://www.bps.go.id>. Data yang digunakan dalam penelitian ini adalah data Banyaknya Desa /Kelurahan Menurut Upaya Antisipasi/Mitigasi Bencana Alam Menurut Provinsi, yang terdiri dari 34 provinsi. Ada 4 variabel yang digunakan yaitu Sistem Peringatan Dini Bencana Alam, Sistem Peringatan Dini Tsunami, Perlengkapan Keselamatan, Jalur Evakuasi. Data akan diolah dengan melakukan clustering Banyaknya Desa/Kelurahan Menurut Upaya Antisipasi / Mitigasi Bencana Alam Menurut Provinsi dalam 3 cluster yakni cluster tingkat antisipasi tinggi, cluster tingkat antisipasi sedang dan cluster tingkat antisipasi rendah.

3.2. Tahap Pengolahan Data

Data yang telah diperoleh akan diolah terlebih dahulu untuk dapat diclustering. Dalam tahap sebelumnya, data setiap provinsi akan dijumlah setiap aspeknya sehingga pada tahapan ini sudah diperoleh perhitungan nilai yang akan diproses pada tahap clustering.

3.3. Tahap Clustering

Clustering merupakan klasifikasi tanpa pengawasan dan merupakan proses partisi sekumpulan objek data dari satu set menjadi beberapa kelas. Hal ini dapat dilakukan dengan menerapkan berbagai persamaan dan langkah-langkah mengenai jarak algoritma, yaitu dengan Euclidean Distance [6]. Analisis kluster ialah algoritma yang dipakai untuk membagi rangkaian data menjadi beberapa grup berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya [7]. Dalam menentukan cluster berdasarkan data yang telah tersedia, dibutuhkan sebuah flowchart untuk memudahkan dalam menentukan alur perhitungan sebagai alur untuk menemukan hasil dari penerapan cluster terhadap data yang akan diproses. Berikut adalah flowchart dalam menentukan cluster dengan K-Means [8].



Gambar 1. Tahapan Clustering

3.4. Tahap Analisis

Pada tahapan ini dilakukan analisis data banyaknya desa/kelurahan menurut upaya antisipasi / mitigasi bencana alam. Data yang diperoleh diolah dengan menggunakan perhitungan bobot dari tiap indeks. Pada tahapan sebelumnya, telah ditentukan akan dicluster ke dalam 3 cluster yakni cluster tingkat antisipasi tinggi, cluster tingkat antisipasi sedang dan cluster tingkat antisipasi rendah. Pada tahapan inilah akan dianalisis hasilnya.

Dalam melakukan clustering, data yang diperoleh akan dihitung terlebih dahulu berdasarkan banyaknya desa/kelurahan menurut upaya antisipasi/mitigasi bencana alam. Adapun kriteria yang digunakan yaitu sebanyak 4 kriteria penilaian seperti yang ditunjukkan pada tabel 1.

Tabel 1 Data Jumlah Desa/Kelurahan Yang Memiliki Sarana Kesehatan Menurut Provinsi

No	Provinsi	Sistem Peringatan Dini Bencana Alam	Sistem Peringatan Dini Tsunami	Perlengkapan Keselamatan	Jalur Evakuasi
1	Aceh	208	67	46	511
2	Sumatera Utara	185	27	59	198
3	Sumatera Barat	186	79	71	263
4	Riau	63	2	62	52
5	Jambi	34	0	22	36
6	Sumatera Selatan	92	1	25	31
7	Bengkulu	108	26	22	236
8	Lampung	171	26	12	120
9	Kepulauan Bangka Belitung	2	1	10	13
10	Kepulauan Riau	10	0	12	17
11	DKI Jakarta	64	0	166	123
12	Jawa Barat	761	44	184	560
13	Jawa Tengah	1,287	40	252	743
14	DI Yogyakarta	191	17	72	148
15	Jawa Timur	904	39	158	624

No	Provinsi	Sistem Peringatan Dini Bencana Alam	Sistem Peringatan Dini Tsunami	Perlengkapan Keselamatan	Jalur Evakuasi
16	Banten	70	12	24	71
17	Bali	511	16	8	38
18	Nusa Tenggara Barat	56	12	16	63
19	Nusa Tenggara Timur	174	9	17	63
20	Kalimantan Barat	19	2	24	49
21	Kalimantan Tengah	26	0	21	69
22	Kalimantan Selatan	26	0	28	51
23	Kalimantan Timur	42	2	49	68
24	Kalimantan Utara	13	0	9	9
25	Sulawesi Utara	266	49	38	242
26	Sulawesi Tengah	100	7	11	54
27	Sulawesi Selatan	100	7	35	61
28	Sulawesi Tenggara	29	2	8	44
29	Gorontalo	31	2	8	14
30	Sulawesi Barat	14	6	5	15
31	Maluku	45	22	9	78
32	Maluku Utara	63	18	35	133
33	Papua Barat	49	16	14	66
34	Papua	42	8	16	48

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

Kemudian data tersebut akan masuk ke tahapan clustering dengan menerapkan algoritma K-Means untuk mengcluster data menjadi tiga cluster.

3.5 Centroid Data

Dalam penerapan algoritma K-means dihasilkan nilai titik tengah atau centroid dari data yang didapat dengan ketentuan bahwa clusterisasi yang diinginkan adalah 3, Penentuan cluster dibagi atas tiga bagian yakni cluster tingkat antisipasi tinggi (C1), cluster tingkat antisipasi sedang (C2) dan cluster tingkat antisipasi rendah (C3). Maka nilai titik tengah atau centroid juga terdapat 3 titik. Penentuan titik cluster ini dilakukan dengan mengambil nilai terbesar (maksimum) untuk cluster tingkat antisipasi tinggi (C1), nilai rata-rata (average) untuk cluster tingkat antisipasi sedang (C2) dan nilai terkecil (minimum) untuk cluster tingkat antisipasi rendah (C3). Nilai titik tersebut dapat diketahui pada Tabel 2 berikut:

Tabel 2. Centroid Data Awal

	A	B	C	D
C1	1,287	79	252	743
C2	175	16	46	144
C3	2	0	5	9

Keterangan :

A = Sistem Peringatan Dini Bencana Alam

B = Sistem Peringatan Dini Tsunami

C = Perlengkapan Keselamatan

D = Jalur Evakuasi

3.6 Clustering Data

Dengan menggunakan centroid tersebut maka dapat dicluster data yang telah didapat menjadi 3 cluster. Proses cluster dengan mengambil jarak terdekat dari setiap data yang diolah dengan menggunakan rumus berikut:

$$d(x,y)=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\dots+(x_m-y_m)^2}$$

Keterangan :

d(x,y) adalah Euclidean Distance

(x) merupakan koordinat object dan

(y) merupakan koordinat centroid.

. Dari data desa/kelurahan menurut upaya antisipasi/mitigasi bencana alam menurut provinsi didapatkan pengelompokan pada iterasi 1 untuk 3 cluster tersebut. Cluster tingkat antisipasi tinggi (C1) yakni Jawa Barat, Jawa Tengah, Jawa Timur cluster tingkat sarana kesehatan sedang (C2) yakni 12 Provinsi dan cluster tingkat sarana

kesehatan rendah (C3) yakni 19 provinsi lainnya . Proses pencarian jarak terpendek, pengelompokan data pada iterasi 1 dan Clustering atas dapat digambarkan pada tabel dan gambar berikut:

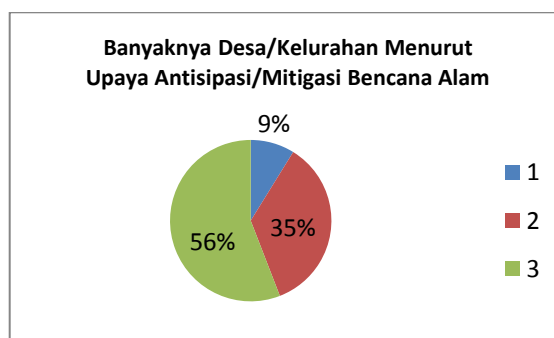
Tabel 3. Perhitungan Jarak Pusat Cluster

Provinsi	Sistem Peringatan Dini Bencana Alam	Sistem Peringatan Dini Tsunami	Perlengkapan Keselamatan	Jalur Evakuasi	C1	C2	C3	Jarak Terpendek
Aceh	208	67	46	511	1122.78	371.52	548.28	371.52
Sumatera Utara	185	27	59	198	1245.54	57.15	269.92	57.15
Sumatera Barat	186	79	71	263	1214.64	136.91	330.10	136.91
Riau	63	2	62	52	1420.45	146.69	93.93	93.93
Jambi	34	0	22	36	1459.11	179.99	45.19	45.19
Sumatera Selatan	92	1	25	31	1411.59	142.75	94.79	94.79
Bengkulu	108	26	22	236	1304.91	116.13	252.45	116.13
Lampung	171	26	12	120	1301.54	42.75	203.98	42.75
Kepulauan Bangka Belitung	2	1	10	13	1499.59	220.51	6.48	6.48
Kepulauan Riau	10	0	12	17	1490.52	211.62	13.30	13.30
DKI Jakarta	64	0	166	123	1376.14	165.87	206.79	165.87
Jawa Barat	761	44	184	560	562.15	732.32	955.86	562.15
Jawa Tengah	1.287	40	252	743	39.00	1280.05	1500.86	39.00
DI Yogyakarta	191	17	72	148	1261.54	31.26	244.58	31.26
Jawa Timur	904	39	158	624	413.87	880.30	1103.07	413.87
Banten	70	12	24	71	1410.37	129.82	94.73	94.73
Bali	511	16	8	38	1078.29	354.67	510.09	354.67
Nusa Tenggara Barat	56	12	16	63	1427.57	147.07	78.08	78.08
Nusa Tenggara Timur	174	9	17	63	1327.14	86.62	180.90	86.62
Kalimantan Barat	19	2	24	49	1465.39	184.51	47.48	47.48
Kalimantan Tengah	26	0	21	69	1450.52	169.39	66.57	66.57
Kalimantan Selatan	26	0	28	51	1457.88	177.31	53.56	53.56
Kalimantan Timur	42	2	49	68	1432.76	153.92	83.79	83.79
Kalimantan Utara	13	0	9	9	1492.35	214.75	11.70	11.70
Sulawesi Utara	266	49	38	242	1157.64	137.69	357.04	137.69
Sulawesi Tengah	100	7	11	54	1395.33	122.68	108.23	108.23
Sulawesi Selatan	100	7	35	61	1387.94	112.93	115.14	112.93
Sulawesi Tenggara	29	2	8	44	1461.72	181.53	44.35	44.35
Gorontalo	31	2	8	14	1474.60	198.24	29.65	29.65
Sulawesi Barat	14	6	5	15	1488.91	210.60	14.70	14.70
Maluku	45	22	9	78	1430.76	150.39	84.32	84.32
Maluku Utara	63	18	35	133	1386.03	112.85	142.55	112.85
Papua Barat	49	16	14	66	1432.34	151.54	76.12	76.12
Papua	42	8	16	48	1446.99	166.94	57.50	57.50

Tabel 4. Hasil Pengelompokan

No	Provinsi	C1	C2	C3
1	Aceh		1	
2	Sumatera Utara		1	
3	Sumatera Barat		1	
4	Riau			1
5	Jambi			1
6	Sumatera Selatan			1
7	Bengkulu		1	
8	Lampung		1	
9	Kepulauan Bangka Belitung			1
10	Kepulauan Riau			1
11	DKI Jakarta		1	
12	Jawa Barat	1		
13	Jawa Tengah	1		
14	DI Yogyakarta		1	
15	Jawa Timur	1		
16	Banten			1
17	Bali		1	
18	Nusa Tenggara Barat			1
19	Nusa Tenggara Timur		1	
20	Kalimantan Barat			1
21	Kalimantan Tengah			1
22	Kalimantan Selatan			1
23	Kalimantan Timur			1
24	Kalimantan Utara			1
25	Sulawesi Utara		1	
26	Sulawesi Tengah			1
27	Sulawesi Selatan		1	
28	Sulawesi Tenggara			1
29	Gorontalo			1
30	Sulawesi Barat			1
31	Maluku			1
32	Maluku Utara		1	

33	Papua Barat			1
34	Papua			1



Gambar 2. Iterasi 1

Proses K-Means akan terus beriterasi sampai pengelompokan data sama dengan pengelompokan data iterasi sebelumnya. Dengan kata lain, proses akan terus melakukan iterasi sampai data pada iterasi terakhir sama dengan iterasi sebelumnya. Pada iterasi 1 diperoleh cluster data banyaknya desa/kelurahan menurut upaya antisipasi/mitigasi bencana alam menurut provinsi, yang dapat dilihat pada gambar 2. Proses iterasi tersebut berhenti pada iterasi ke 3, pada iterasi 2 akan dilakukan proses mencari nilai titik tengah atau centroid yang dapat dengan menggunakan rumus berikut :

$$\bar{v}_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} x_{kj}$$

Keterangan :

v_{ij} adalah centroid/ rata-rata cluster ke-I untuk variable ke-j

N_i adalah jumlah data yang menjadi anggota cluster ke-i

i, k adalah indeks dari cluster

j adalah indeks dari variabel

x_{kj} adalah nilai data ke-k yang ada di dalam cluster tersebut untuk variable ke-j

Hasilnya dapat diketahui pada Tabel 6 berikut :

Tabel 6. Data Centroid Iterasi 4

	A	B	C	D
C1	984	41	198	642
C2	222.22	35.11	38.33	202.11
C3	45.00	5.45	27.68	52.95

Keterangan :

A = Sistem Peringatan Dini Bencana Alam

B = Sistem Peringatan Dini Tsunami

C = Perlengkapan Keselamatan

D = Jalur Evakuasi

Setelah mendapatkan nilai titik tengah atau centroid, proses sama dilakukan dengan mencari jarak terdekat. Proses pencarian jarak terdekat, pengelompokan data pada iterasi 3 dan Clustering data dapat digambarkan Pada tabel dan gambar berikut:

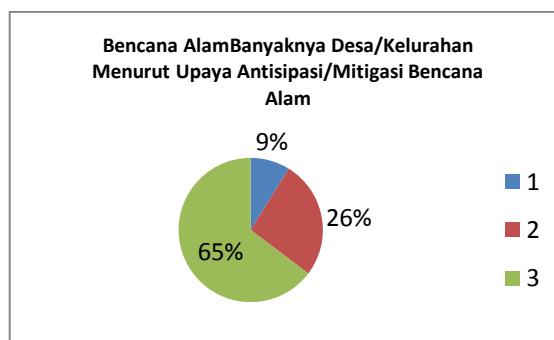
Tabel 7 Perhitungan Jarak Pusat Cluster Iterasi 3

Provinsi	Sistem Peringatan Dini Bencana Alam	Sistem Peringatan Dini Tsunami	Perlengkapan Keselamatan	Jalur Evakuasi	C1	C2	C3	Jarak Terpendek
Aceh	208	67	46	511	802.00	310.95	490.41	310.95
Sumatera Utara	185	27	59	198	924.85	43.54	205.14	43.54
Sumatera Barat	186	79	71	263	893.46	89.51	266.99	89.51
Riau	63	2	62	52	1103.06	222.58	38.92	38.92
Jambi	34	0	22	36	1141.40	254.01	21.69	21.69
Sumatera Selatan	92	1	25	31	1095.87	218.12	52.13	52.13
Bengkulu	108	26	22	236	981.67	120.60	194.75	120.60
Lampung	171	26	12	120	984.19	100.71	145.05	100.71
Kepulauan Bangka Belitung	2	1	10	13	1182.09	293.64	61.46	61.46
Kepulauan Riau	10	0	12	17	1173.03	285.01	52.85	52.85

Provinsi	Sistem Peringatan Dini Bencana Alam	Sistem Peringatan Dini Tsunami	Perlengkapan Keselamatan	Jalur Evakuasi	C1	C2	C3	Jarak Terpendek
DKI Jakarta	64	0	166	123	1057.74	220.96	156.30	156.30
Jawa Barat	761	44	184	560	238.14	663.07	892.00	238.14
Jawa Tengah	1,287	40	252	743	323.82	1213.26	1438.83	323.82
DI Yogyakarta	191	17	72	148	943.22	73.24	180.13	73.24
Jawa Timur	904	39	158	624	91.32	810.65	1040.23	91.32
Banten	70	12	24	71	1092.22	202.73	31.73	31.73
Bali	511	16	8	38	790.99	334.08	466.77	334.08
Nusa Tenggara Barat	56	12	16	63	1109.40	219.12	20.03	20.03
Nusa Tenggara Timur	174	9	17	63	1012.68	151.04	129.88	129.88
Kalimantan Barat	19	2	24	49	1146.76	256.99	26.78	26.78
Kalimantan Tengah	26	0	21	69	1131.14	240.32	26.32	26.32
Kalimantan Selatan	26	0	28	51	1139.31	250.35	19.87	19.87
Kalimantan Timur	42	2	49	68	1113.98	227.32	26.49	26.49
Kalimantan Utara	13	0	9	9	1175.31	288.37	57.75	57.75
Sulawesi Utara	266	49	38	242	837.53	60.83	294.25	60.83
Sulawesi Tengah	100	7	11	54	1078.76	195.99	57.50	57.50
Sulawesi Selatan	100	7	35	61	1071.04	188.82	56.09	56.09
Sulawesi Tenggara	29	2	8	44	1143.52	253.67	27.12	27.12
Gorontalo	31	2	8	14	1157.86	271.97	45.96	45.96
Sulawesi Barat	14	6	5	15	1171.72	283.42	54.00	54.00
Maluku	45	22	9	78	1111.88	218.73	35.36	35.36
Maluku Utara	63	18	35	133	1065.25	174.45	83.32	83.32
Papua Barat	49	16	14	66	1113.94	222.46	22.01	22.01
Papua	42	8	16	48	1129.07	239.72	13.29	13.29

Tabel 8. Hasil Pengelompokan

No	Provinsi	C1	C2	C3
1	Aceh		1	
2	Sumatera Utara		1	
3	Sumatera Barat		1	
4	Riau			1
5	Jambi			1
6	Sumatera Selatan			1
7	Bengkulu		1	
8	Lampung		1	
9	Kepulauan Bangka Belitung			1
10	Kepulauan Riau			1
11	DKI Jakarta			1
12	Jawa Barat	1		
13	Jawa Tengah	1		
14	DI Yogyakarta		1	
15	Jawa Timur	1		
16	Banten			1
17	Bali		1	
18	Nusa Tenggara Barat			1
19	Nusa Tenggara Timur		1	
20	Kalimantan Barat			1
21	Kalimantan Tengah			1
22	Kalimantan Selatan			1
23	Kalimantan Timur			1
24	Kalimantan Utara			1
25	Sulawesi Utara		1	
26	Sulawesi Tengah			1
27	Sulawesi Selatan			1
28	Sulawesi Tenggara			1
29	Gorontalo			1
30	Sulawesi Barat			1
31	Maluku			1
32	Maluku Utara			1
33	Papua Barat			1
34	Papua			1



Gambar 3. Iterasi 2

3.7 Analisa Data

Pada iterasi 3, pengelompokan data yang dilakukan terhadap 3 cluster dengan iterasi 2 didapatkan hasil yang sama. Dari 34 data jumlah desa/kelurahan yang memiliki sarana kesehatan menurut provinsi dapat diketahui, 3 provinsi cluster tingkat tinggi yakni Jawa Barat, Jawa Tengah, Jawa Timur, 9 provinsi cluster tingkat sedang dan 22 provinsi lainnya termasuk cluster tingkat rendah.

4. KESIMPULAN

Dari permasalahan diatas tentang data banyaknya desa/kelurahan menurut upaya antisipasi/mitigasi bencana alam menurut provinsi dapat diselesaikan dengan teknik data mining menggunakan aturan clustering untuk mengelompokkan desa/kelurahan menurut upaya antisipasi/mitigasi bencana alam di Indonesia. Data tersebut diolah menggunakan Ms. Excel untuk ditentukan nilai centroid dalam 3 cluster yaitu cluster tingkat tinggi, cluster tingkat sedang dan cluster tingkat rendah. Sehingga diperoleh penilaian berdasarkan indeks indeks desa/kelurahan menurut upaya antisipasi/mitigasi bencana alam dengan 3 provinsi tingkat tinggi yaitu Jawa Barat, Jawa Tengah, Jawa Timur, 9 provinsi tingkat sedang, dan 22 provinsi lainnya termasuk tingkat rendah. Hasil yang didapat dari penelitian dapat menjadi masukan kepada pemerintah, provinsi yang menjadi perhatian lebih dalam peningkatan sarana dan prasarana pada desa/kelurahan menurut upaya antisipasi/mitigasi bencana alam berdasarkan cluster yang telah dilakukan.

5. REFERENSI

- [1] A. P. Windarto, "Implementation of Data Mining on Rice Imports by Major Country of Origin Using Algorithm Using K-Means Clustering Method," *Int. J. Artif. Intell. Res.*, vol. 1, no. 2, pp. 26–33, 2017.
- [2] U. R. Raval and C. Jani, "Implementing and Improvisation of K-means Clustering," *Int. J. Comput. Sci. Mob. Comput.*, vol. 5, no. 5, pp. 72–76, 2016.
- [3] M. K. Arzoo, A. Prof, and K. Rathod, "K-Means algorithm with different distance metrics in spatial data mining with uses of NetBeans IDE 8 . 2," *Int. Res. J. Eng. Technol.*, vol. 4, no. 4, pp. 2363–2368, 2017.
- [4] S. Kumar and S. K. Rathi, "Performance Evaluation of K-Means Algorithm and Enhanced Mid-point based K-Means Algorithm on Mining Frequent Patterns," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 4, no. 10, pp. 545–548, 2014.
- [5] A. Yadav and S. Dhingra, "An Enhanced K-Means Clustering Algorithm to Remove Empty Clusters," *IJEDR*, vol. 4, no. 4, pp. 901–907, 2016.
- [6] A. P. Windarto, L. S. Dewi, and D. Hartama, "Implementation of Artificial Intelligence in Predicting the Value of Indonesian Oil and Gas Exports With BP Algorithm," *Int. J. Recent Trends Eng. Res.*, vol. 3, no. 10, pp. 1–12, 2017.
- [7] Sumijan, A. P. Windarto, A. Muhammad, and Budiharjo, "Implementation of Neural Networks in Predicting the Understanding Level of Students Subject," *Int. J. Softw. Eng. Its Appl.*, vol. 10, no. 10, pp. 189–204, 2016.
- [8] A. Wanto and A. P. Windarto, "Analisis Prediksi Indeks Harga Konsumen Berdasarkan Kelompok Kesehatan Dengan Menggunakan Metode Backpropagation," *J. Penelit. Tek. Inform.*, vol. 2, no. 2, pp. 37–44, 2017.
- [9] A. Wanto, A. P. Windarto, D. Hartama, and I. Parlina, "Use of Binary Sigmoid Function And Linear Identity In Artificial Neural Networks For Forecasting Population Density," *Int. J. Inf. Syst. Technol.*, vol. 1, no. 1, pp. 43–54, 2017.
- [10] A. N. D. J. D. Fadhilah, "Perancangan Aplikasi Sistem Pakar PEnyakit Kulit Pada Anak Dengan Metode Expert System Development Life Cycle," *J. Algoritm. Sekol. Tinggi Teknol. Garut*, vol. 9, no. 13, pp. 1–7, 2012.
- [11] S. Fekri-Ershad, H. Tajalizadeh, and S. Jafari, "Design and Development of an Expert System to Help Head of University Departments," *Int. J. Sci. Mod. Eng.*, vol. 1, no. 2, pp. 45–48, 2013.

- [12] M. Min, "A rule based expert system for analysis of mobile sales data on fashion market," *2013 Int. Conf. Inf. Sci. Appl. ICISA 2013*, 2013.
- [13] M. Mohammadi and S. Jafari, "An expert system for recommending suitable ornamental fish addition to an aquarium based on aquarium condition," *arXiv Prepr. arXiv1405.1524*, vol. 3, no. 2, pp. 1–7, 2014.
- [14] I. Chen and B. L. Poole, "Performance Evaluation of Rule Grouping on a Real-Time Expert System Architecture," vol. 6, no. 6, pp. 883–891, 2014.