

Analisa Metode K-Means pada Pengelompokan Kriminalitas Menurut Wilayah

Sinta Maulina Dewi, Agus Perdana Windarto, Irfan Sudahri Damanik, Heru Satria

Sistem Informasi, STIKOM Tunas Bangsa Pematangsiantar

Email: ¹sintamaulinadewi033@gmail.com, ²agus.perdana@amiktunasbangsa.ac.id

Abstrak

Banyaknya tindak kejahatan saat ini yang menjadi perhatian dan pemberitaan di berbagai media massa maupun elektronik. Aksi kejahatan yang sering terjadi ialah pembunuhan, baik pembunuhan terencana maupun tidak. Aksi ini banyak menimbulkan kekhawatiran terhadap masyarakat dikarenakan sudah banyaknya korban yang ada disekitar mereka. Penelitian ini membahas tentang AnalisaMetode K-means pada Pengelompokan Kriminalitas Menurut Provinsi. Sumber data pada penelitian ini merupakan kumpulan dari berbagai dokumen-dokumen keterangan Aksi Kriminalitas oleh Badan Pusat Statistik Nasional. Data yang digunakan pada penelitian ini adalah data dari tahun 2012 yang terdiri dari 31 provinsi. Variable yang digunakan yaitu cluster tingkat tinggi dan cluster tingkat rendah.

Kata Kunci: Kriminalitas, Sistem Pakar, K-Means, Clustering

1. PENDAHULUAN

Kriminal merupakan segala macam bentuk tindakan dan perbuatan yang merugikan secara ekonomis dan psikologis yang melanggar hukum yang berlaku dalam Negara Indonesia serta norma-norma social dan agama[1]. Tindak kejahatan yang dianggap kriminalitas yaitu perampokan, pemerkosaan, bahkan hingga pembunuhan. Banyaknya tindak kejahatan saat ini yang menjadi perhatian dan pemberitaan di berbagai media massa maupun elektronik. Aksi kejahatan yang sering terjadi ialah pembunuhan, baik pembunuhan terencana maupun tidak. Aksi ini banyak menimbulkan kekhawatiran terhadap masyarakat.

Analisa metode k-means merupakan metode clustering yang paling rendah dan sederhana, hal ini dikarenakan k-means mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien[2]. Penggunaan metode ini menjadi alternatif disebabkan adanya penelitian sebelumnya[3]. Dalam metode ini perlu adanya pencarian pusat cluster yang ditentukan pada jarak minimum setiap data pada cluster[4]. Data yang digunakan berdasarkan kumpulan dokumen yang dihasilkan oleh Badan Pusat Statistik Nasional. Hasil dari penelitian ini menjadi masukan bagi pemerintah agar provinsi yang mengalami tindak kriminalitas cluster tinggi mendapat perhatian lebih. Proses analisa melalui pembagian cluster yang menjadi 2 bagian diantaranya cluster rendah dan cluster tinggi.

2. METODOLOGI PENELITIAN

Dalam penelitian ini menggunakan metode datamining antara lain: a) Tahap pengumpulan data, b) Tahap pengolahan data, c) Tahap clustering dan d) Tahap Analisis

A. Tahap Pengumpulan Data

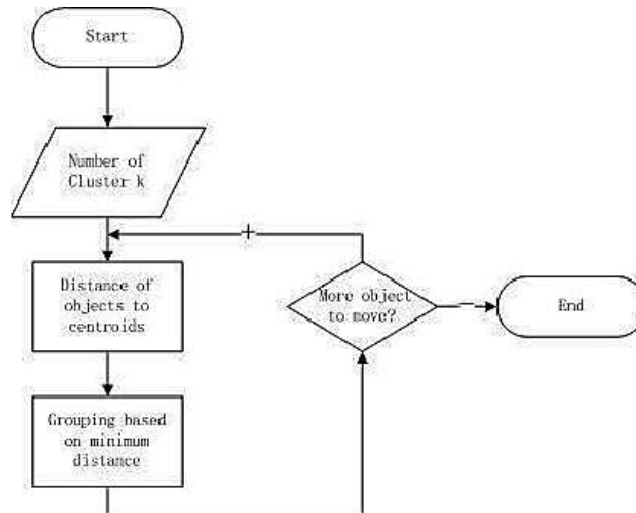
Dalam Analisa Data Mining K-Means[5]Pada Pengelompokan Kriminalitas Menurut Provinsi, diperlukan data yang terkait tentang hal tersebut. Sumber data penelitian diperoleh dari data yang dikumpulkan berdasarkan kumpulan dokumen-dokumen kriminalitas pada tahun 2014.

B. Tahap Pengolaan Data

Data yang telah diperoleh kemudian diolah terlebih dahulu untuk dapat dicluster. Sebelum itu, data setiap provinsi dijumlah setiap aspeknya sehingga pada tahap ini sudah diperoleh perhitungan nilai yang akan diproses.

C. Tahap Clustering

Clustering merupakan pengelompokkan sejumlah data atau objek kedalam cluster[2]. Dalam menentukan suatu cluster berdasarkan data yang ada, dibutuhkan adanya flowchart dalam memudahkan penentuan alur perhitungan untuk mendapatkan hasil dari penerapan cluster terhadap data yang diproses. Berikut adalah flowchart dalam menentukan cluster k-means :



Gambar 1. Diagram flowchart

D. Tahap Analisis

Dalam menganalisa pengelompokkan kriminalitas. Data yang telah diperoleh, kemudian diolah menggunakan perhitungan bobot dari tiap indeks. Pada tahapan sebelumnya, telah ditentukan akan dicluster ke dalam 2 cluster diantaranya cluster rendah dan cluster tinggi, yang hasilnya akan dianalisis.

3. ANALISA DAN PEMBAHASAN

Dalam melakukan clustering, data yang telah diperoleh akan dihitung berdasarkan jumlah data kriminalitas pada tahun 2012 menurut provinsi. Berikut data kriminalitas pada tahun 2012 :

Tabel 1. Data kriminalitas menurut provinsi

No	Kepolisian Daerah	x
1	Aceh	44
2	Sumatera Utara	143
3	Sumatera Barat	14
4	Riau	30
5	Jambi	26
6	Sumatera Selatan	129
7	Bengkulu	27
8	Lampung	25
9	Bangka Belitung	28
10	Kepulauan Riau	12
11	Metro Jaya	72
12	Jawa Barat	107
13	Jawa Tengah	48
14	DI Yogyakarta	17
15	Jawa Timur	66
16	Banten	19
17	Bali	14
18	Nusa Tenggara Barat	24
19	Nusa Tenggara Timur	53
20	Kalimantan Barat	15
21	Kalimantan Tengah	36
22	Kalimantan Selatan	31
23	Kalimantan Timur	61
24	Sulawesi Utara	49
25	Sulawesi Tengah	84
26	Sulawesi Selatan	115
27	Sulawesi Tenggara	45
28	Gorontalo	8

No	Kepolisian Daerah	x
29	Maluku	12
30	Maluku Utara	4
31	Papua	98

Data tersebut kemudian memasuki tahap analisa, data akan dicluster menjadi 2 cluster diantaranya cluster tinggi dan cluster rendah.

a. *Centroid Data*

Dalam menerapkan algoritma K-means akan menghasilkan centroid atau titik tengah cluster dibagi menjadi 2 bagian, maka titik tengah atau centroid terdapat 2 titik diantaranya, (C1) cluster tinggi dan (C2) cluster rendah. Pada titik ini dilakukan dengan nilai tinggi (maximum), dan nilai rendah (minimum). Nilai titik tersebut dapat dilihat pada table berikut:

Tabel 2. Centroid data awal

Atribut	C1	C2
Nilai	143	4

b. *Clustering Data*

Dari centroid tersebut, maka didapat 2 cluster dengan mengambil jarak terdekat dari setiap data yang diolah. Dengan didapatkan pengelompokkan pada iterasi 1 untuk 2 cluster tersebut. Cluster tinggi (C1) diantaranya Sumatera Utara, Sumatera Selatan, Jawa Barat, Sulawesi Tengah, Sulawesi Selatan, dan Papua. Dan cluster rendah (C2) diantaranya Aceh, Sumatera Barat, Riau, Jambi, Bengkulu, dan 20 provinsi lainnya. Pada proses tersebut dapat dilihat pada table berikut :

Tabel 3. Perhitungan jarak pusat cluster

No	Kepolisian Daerah	Jlh Kriminalitas	C1	C2	Jarak Terpendek
1	Aceh	44	99	40	40
2	Sumatera Utara	143	0	139	0
3	Sumatera Barat	14	129	10	10
4	Riau	30	113	26	26
5	Jambi	26	117	22	22
6	Sumatera Selatan	129	14	125	125
7	Bengkulu	27	116	23	23
8	Lampung	25	118	21	21
9	Bangka Belitung	28	115	24	24
10	Kepulauan Riau	12	131	8	8
11	Metro Jaya	72	71	68	68
12	Jawa Barat	107	36	103	103
13	Jawa Tengah	48	95	44	44
14	DI Yogyakarta	17	126	13	13
15	Jawa Timur	66	77	62	62
16	Banten	19	124	15	15
17	Bali	14	129	10	10
18	Nusa Tenggara Barat	24	119	20	20
19	Nusa Tenggara Timur	53	90	49	49
20	Kalimantan Barat	15	128	11	11
21	Kalimantan Tengah	36	107	32	32
22	Kalimantan Selatan	31	112	27	27
23	Kalimantan Timur	61	82	57	57
24	Sulawesi Utara	49	94	45	45
25	Sulawesi Tengah	84	59	80	59
26	Sulawesi Selatan	115	28	111	111
27	Sulawesi Tenggara	45	98	41	41
28	Gorontalo	8	135	4	4
29	Maluku	12	131	8	8
30	Maluku Utara	4	139	0	0
31	Papua	98	45	94	94

Setelah proses diatas, maka didapatkan hasil sebagai berikut :

Tabel 4. Pengelompokan dari iterasi 1

No	Kepolisian Daerah	C1	C2
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	Bangka Belitung		1
10	Kepulauan Riau		1
11	Metro Jaya		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	Sulawesi Utara		1
25	Sulawesi Tengah	1	
26	Sulawesi Selatan	1	
27	Sulawesi Tenggara		1
28	Gorontalo		1
29	Maluku		1
30	Maluku Utara		1
31	Papua	1	

Proses K-means akan terus beriterasi sampai pengelompokkan sata sama dengan pengelompokkan data sebelumnya. Maka proses tersebut akan terus iterasi sampai data pada iterasi terakhir sama dengan iterasi sebelumnya. Proses iterasi tersebut berhenti pada iterasi ke 5, pada iterasi ini akan dilakukan pencarian nilai titik tengah atau centroid pada table berikut :

Tabel 6. Data centroid iterasi 5

Atribut	C1	C2
Nilai	101.75	27.91304348

Setelah mendapatkan nilai titik tengah atau centroid, proses sama dilakukan dengan mencari jarak terdekat. Proses pencarian jarak terpendek, pengelompokan data pada iterasi 5 dan Clustering data dapat digambarkan pada tabel berikut:

Table 7. perhitungan jarak pusat cluster iterasi 5

No	Kepolisian Daerah	Jlh Kriminalitas	C1	C2	Iterasi5
1	Aceh	44	57.75	16.0870	16.087
2	Sumatera Utara	143	41.25	115.0870	41.25
3	Sumatera Barat	14	87.75	13.9130	13.913
4	Riau	30	71.75	2.0870	2.087
5	Jambi	26	75.75	1.9130	1.913



No	Kepolisian Daerah	Jlh Kriminalitas	C1	C2	Iterasi5
6	Sumatera Selatan	129	27.25	101.0870	27.25
7	Bengkulu	27	74.75	0.9130	75.75
8	Lampung	25	76.75	2.9130	76.75
9	Bangka Belitung	28	73.75	0.0870	73.75
10	Kepulauan Riau	12	89.75	15.9130	15.913
11	Metro Jaya	72	29.75	44.0870	29.75
12	Jawa Barat	107	5.25	79.0870	5.25
13	Jawa Tengah	48	53.75	20.0870	20.087
14	DI Yogyakarta	17	84.75	10.9130	10.913
15	Jawa Timur	66	35.75	38.0870	35.75
16	Banten	19	82.75	8.9130	8.913
17	Bali	14	87.75	13.9130	13.913
18	Nusa Tenggara Barat	24	77.75	3.9130	3.913
19	Nusa Tenggara Timur	53	48.75	25.0870	25.087
20	Kalimantan Barat	15	86.75	12.9130	12.913
21	Kalimantan Tengah	36	65.75	8.0870	8.087
22	Kalimantan Selatan	31	70.75	3.0870	3.087
23	Kalimantan Timur	61	40.75	33.0870	33.087
24	Sulawesi Utara	49	52.75	21.0870	21.087
25	Sulawesi Tengah	84	17.75	56.0870	17.75
26	Sulawesi Selatan	115	13.25	87.0870	13.25
27	Sulawesi Tenggara	45	56.75	17.0870	17.087
28	Gorontalo	8	93.75	19.9130	19.913
29	Maluku	12	89.75	15.9130	15.913
30	Maluku Utara	4	97.75	23.9130	23.913
31	Papua	98	3.75	70.0870	3.75

Tabel 8. Pengelompokkan data iterasi 5

No	Kepolisian Daerah	C1	C2
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	Bangka Belitung		1
10	Kepulauan Riau		1
11	Metro Jaya	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	Sulawesi Utara		1
25	Sulawesi Tengah	1	

No	Kepolisian Daerah	C1	C2
26	Sulawesi Selatan	1	
27	Sulawesi Tenggara		1
28	Gorontalo		1
29	Maluku		1
30	Maluku Utara		1
31	Papua	1	

c. **Analisi Data**

Dari data diatas didapatkan hasil yang sama dengan iterasi sebelumnya, maka hasil dari iterasi ke 5 dapat diketahui bahwa, cluster tinggi (C1) diantaranya Sumatera Utara, Sumater Selatan, Metro Jaya, Jawa Barat, Jawa Timur, Sulawesi Selatan, Sulawesi Tengah, dan Papua. Dan cluster rendah (C2) diantaranya Aceh, Sumatera Barat, Riau, Jambi, Bengkulu, dan 18 provinsi lainnya.

4. KESIMPULAN

Dari penelitian diatas dapat disimpulkan bahwa, Analisa Metode K-Means Pada Pengelompokkan Kriminalitas Menurut Provinsi didapatkan hasil bahwa cluster tinggi (C1) = 8 provinsi, diantaranya Sumatera Utara, Sumater Selatan, Metro Jaya, Jawa Barat, Jawa Timur, Sulawesi Selatan, Sulawesi Tengah, dan Papua. Dan cluster rendah (C2) = 23 Provinsi diantaranya Aceh, Sumatera Barat, Riau, Jambi, Bengkulu, dan 18 provinsi lainnya.

REFERENCES

- [1] D. Hendri, "Kriminalitas :," pp. 1–19, 2014.
- [2] T. Alfina, B. Santosa, and A. R. Barakbah, "145483-ID-analisa-perbandingan-metode-hierarchical," *Tek. ITS Vol.1*, vol. 1, 2012.
- [3] M. G. Sadewo, A. P. Windarto, and D. Hartama, "Penerapan Datamining Pada Populasi Daging Ayam Ras Pedaging Di Indonesia Berdasarkan Provinsi Menggunakan K-Means Clustering," *InfoTekJar (Jurnal Nas. Inform. dan Teknol. Jaringan)*, vol. 2, no. 1, pp. 60–67, 2018.
- [4] K. Fatmawati and A. P. Windarto, "DATA MINING : PENERAPAN RAPIDMINER DENGAN K-MEANS CLUSTER PADA DAERAH TERJANGKIT DEMAM BERDARAH DENGUE (DBD) BERDASARKAN PROVINSI," vol. 3, no. 2, pp. 173–178, 2018.
- [5] 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.