

## ANALISIS ALGORITMA K-MEANS UNTUK KLASTERISASI TINDAK PIDANA KORUPSI DI WILAYAH HUKUM INDONESIA

Wahyu Saputro<sup>1</sup>, Moh Reza Pahlevi<sup>2</sup>, Arief Wibowo<sup>3</sup>

<sup>1,2,3</sup>Magister Ilmu Komputer, Fakultas Ilmu Komputer  
Universitas Budi Luhur

Email: <sup>1</sup>wahyu.saputro6982@gmail.com, <sup>2</sup>onowae848@gmail.com, <sup>3</sup>arief.wibowo@budiluhur.ac.id

(Naskah masuk : 24 Juli 2020, diterima untuk diterbitkan: 23 September 2020)

### Abstrak

Pasca kemerdekaan, Indonesia masih menghadapi perang yaitu perang melawan korupsi yang sudah menjadi budaya di Indonesia selama ini. Ditahun 1999 pemerintah telah mengeluarkan UU No 31 tentang pemberantasan korupsi dan pada tahun 2001 pemerintah memperbaruinya dengan UU No 20. Menurut UU No. 31 Thn 1999 tentang tindak pidana korupsi dan sudah dilakukan pembaruan dengan UU No. 20 Thn 2001, definisi korupsi atau rasuah merupakan sebuah perbuatan tidak jujur atau penyelewengan yang dilakukan karena adanya suatu pemberian. Dalam prakteknya, korupsi lebih dikenal sebagai menerima uang yang ada hubungannya dengan jabatan atas imbalan tertentu yang diberikan. Menurut data dari Kejaksaan RI tingkat korupsi di Indonesia makin tahun makin meningkat baik dari tingkat daerah hingga tingkat Pusat. Peneletian ini menggunakan metode pengumpulan data dari data simkari Kejaksaan RI. Dan untuk analisis menggunakan *data mining* dengan metode *klustering* dengan *algoritma K-Means*. Hasil penelitian di peroleh 3 cluster/kelompok berdasarkan kelompok tingkat kerawanan tindak pidana korupsi. Hasil evaluasi berupa Davies Bouldin Index (DBI) sudah dituliskan pada bagian ini. 2 klaster 0,316 sementara pada 3 klaster 0,113 dan pemodelan 4 klaster 0,618. Dengan demikian model terbaik adalah pada terbentuknya 3 klaster dengan nilai DBI sebesar 0,113. DBI terkecil menyatakan bahwa model memiliki bentuk klaster terbaik sebanyak 3. Hasil penelitian menunjukkan bahwa wilayah dengan kategori rawan tindak pidana korupsi berada pada pulau sumatra ada 5 wilayah, pulau jawa ada 3 wilayah dan pulau sulawesi dan kalimantan ada 1 wilayah.

**Kata kunci:** Korupsi, K-Means, Klustering, Data mining

## K-MEANS ALGORITHM ANALYSIS FOR CORRUPTION CRIMINAL CLUSTERIZATION IN INDONESIA'S LAW

### Abstract

*Post-independence, Indonesia is still facing a war, namely the war against corruption which has become a culture in Indonesia so far. In 1999 the government issued Law No. 31 on corruption eradication and in 2001 the government changed it with Law No. 20. According to Law no. 31 of 1999 concerning criminal acts of corruption and reformation has been carried out by Law no. 20 Year 2001, the definition of corruption or rasuah is an act of dishonesty or fraud committed because of a gift. In practice, corruption is better known as receiving money in connection with a position for certain rewards given. According to data from the Indonesian Prosecutor's Office, the level of corruption in Indonesia is increasing every year, from the regional to the central level. This research uses the method of collecting data from simkari data from the RI Attorney General's Office. And for analysis using data mining with the clustering method with the K-Means algorithm. The research results were obtained by 3 clusters / groups based on the level of vulnerability of corruption. The results of the evaluation in the form of the Davies Bouldin Index (DBI) have been written in this section. 2 clusters 0.316 while in 3 clusters 0.113 and modeling 4 clusters 0.618. Thus, the best model is the formation of 3 clusters with a DBI value of 0.113. The smallest DBI states that the model has the best cluster shape as many as 3. The results show that the regions with the category of prone to corruption are 5 regions on the island of Sumatra, 3 in Java, and 1 in Sulawesi and Kalimantan.*

**Keywords:** Corruption, K-Means, Clustering, Data mining

## 1. PENDAHULUAN

Pasca kemerdekaan, Indonesia masih menghadapi perang yaitu perang melawan korupsi

yang sudah menjadi budaya di Indonesia selama ini. Ditahun 1999 pemerintah telah mengeluarkan UU No 31 tentang pemberantasan korupsi dan pada tahun 2001 pemerintah memperbaruinya dengan UU No 20, definisi korupsi atau rasuah merupakan sebuah tindakan perbuatan tidak jujur atau penyelewengan yang dilakukan karena adanya suatu pemberian. Dalam prakteknya, korupsi lebih dikenal sebagai menerima uang yang ada hubungannya dengan jabatan atas imbalan tertentu yang diberikan. Adapun rumusan dari jenis tindak korupsi, yang di kelompokkan diantaranya kerugian keuangan negara, penyuapan, penyalahgunaan jabatan, curang, kepentingan dalam pengadaan, gratifikasi.

Dalam upaya memberantas tindak pidana korupsi, pemerintah telah mengeluarkan beberapa regulasi terhadap usaha untuk memberantas korupsi, sebagai contoh peraturan yang mengatur kejahatan korupsi dalam hal ini Kejaksaan RI dan Polri memainkan peranan penting dalam memberantas korupsi. Oleh karena itu perlu adanya terobosan dalam menangani kejahatan korupsi tersebut, yang diantaranya mendorong transparansi penyelenggaraan negara, memberikan rekomendasi kepada kementerian dan lembaga untuk melakukan langkah-langkah perbaikan, memoderinsasi pelayanan publik dengan online atau berbasis IT dan dilakukan pengawasan yang terintegritas.

Studi terdahulu yang melakukan analisis data korupsi telah dilakukan, yaitu Analisis hasil audit pemerintahan dan tingka korupsi pemerintahan provinsi di indonesia dengan teori keagenan opini audit dan temuan audit tidak berpengaruh terhadap korupsi [11]. Penelitian lain juga menganalisis data korupsi dengan metode *Support Vector Machine* yang hasilnya bahwa model SVM dapat melakukan pengklasifikasi respon publik tentang isi RUU KPK dengan baik [2]. Se jauh pengamatan yang dilakukan, masih sangat minim jurnal yang menggunakan data korupsi. Untuk kasus korupsi di Indonesia bahkan belum pernah diklasterisasi dengan metode ilmiah. Penelitian ini mengusulkan metode klasterisasi data mining untuk mengelompokkan wilayah hukum berdasarkan data kasus korupsi dari SIMKARI (Sistem Informasi Manajemen Kejaksaan Republik Indonesia)

Untuk teknik pengolahan data menggunakan teknik data mining, dengan menggunakan Algoritma K-Means, karena K-Means merupakan salah satu algoritma *clustering*. Yang terbukti sangat populer dalam model klasterisasi. Algoritma K-Means merupakan suatu algoritma yang sederhana, mudah untuk di implementasikan dan mampu untuk mengklaster data yang besar [7]. untuk mengklasifikasikan atau mengelompokkan sejumlah besar obyek dengan atribut tertentu kedalam kelompok-kelompok sejumlah K.

Dengan memanfaatkan metode K-Means supaya untuk menganalisa korupsi di Indonesia. Pendekatan dibentuk dengan cara digabungkannya

teknik klasterisasi pada K-Means diperuntukan untuk pengelompokkan data yang sama. Keuntungan atas data-data yang sama sangat signifikan akan timbul jika menggunakan teknik klastering .

Berdasarkan pada latar belakang, terdapat permasalahan yang terjadi yaitu upaya pengelompokan kasus korupsi di wilayah hukum Indonesia. Dalam upaya meminimalisir tindak pidana korupsi diperlukan metode untuk mengelompokkan tindak pidana korupsi tersebut sehingga bisa Kejaksaan atau KPK dalam membasmi dan menangkap para pelaku tindak pidana korupsi. Pada penelitian ini, dirumuskan permasalahan dengan cara diantaranya Menggunakan kedekatan jarak *euclidean distance* untuk menganalisa K-Means *Clustering* dan melakukan pengelompokan data ke dalam kategori tindak pidana korupsi paling tinggi, menengah dan cukup menggunakan metode *clustering*.

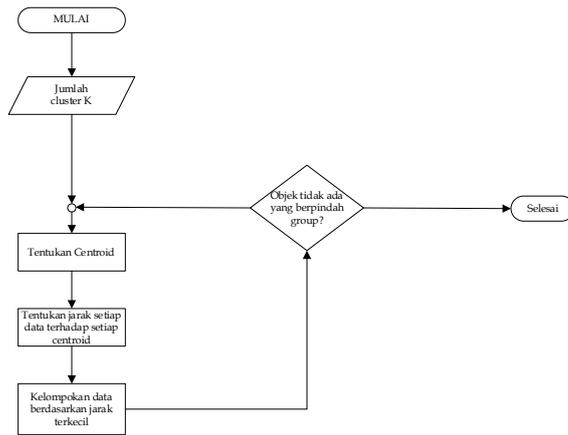
Dalam penelitian ini bertujuan untuk bagaimana metode K-means Klasterisasi bisa digunakan dalam mengelompokkan data-data tentang tindak pidana korupsi diantaranya dapat membantu KPK, Kepolisian dan Kejaksaan dalam mengelompokkan kejahatan korupsi, memberikan gambaran kepada KPK maupun Kejaksaan RI, daerah mana saja yang sering terlibat tindak pidana korupsi, dapat memberikan acuan dalam memetakan daerah rawan tindak pidana korupsi di wilayah Indonesia.

Pada penelitian ini, menggunakan jenis model penelitian eksperimen yaitu mengelompokkan tindak pidana korupsi di Indonesia dengan menggunakan Algoritma K-means sebagai fungsi pengelompokan atau klasterisasi yang sangat besar jumlahnya dan untuk penelitian ini mengambil dataset tindak pidana korupsi pada tahun 2014 dan data diambil dari data dari Kejaksaan RI, *variabel* yang digunakan adalah tahun SPDP, tahun Penuntutan yang ditangani oleh Kejaksaan, dalam penelitian nanti dilakukan penyamaran nama Kejati menjadi nama pulau atau pulau terdekatnya.

## 2. METODE PENELITIAN

Penelitian ini akan menggunakan data yang diolah merupakan dataset tindak pidana korupsi dari tahun 2014. Dataset tindak pidana korupsi terdiri dari atribut Jumlah SPDP, Jumlah Penuntutan, Belum Selesai, berdasarkan wilayah. Data yang akan diuji terdiri dari 1102 data tindak pidana korupsi seluruh Indonesia

Pada penelitian ini pengelompokan tindak pidana korupsi dari data yang diperoleh menggunakan algoritma K-Means. Karena K-means merupakan cara *clustering* yang paling simpel. Karena K-means memiliki kemampuan untuk menggabungkan data yang cukup besar jumlah dalam waktu yang relatif cepat dan efektif. Metode K-Means adalah suatu pola yang menggunakan algoritma klasterisasi data yang berbasis jarak untuk dibagi ke dalam *cluster* dan untuk atributnya hanya berbentuk numerik.



Gambar 2. Flowchart Algoritma K-Means Klasterisasi

Langkah – langkah penelitian untuk pengambilan sampel yang diantaranya sebagai berikut:

- Tentukan K sebagai jumlah cluster yang akan dibentuk
- Buat K centroid secara random
- Hitung jarak setiap data ke masing – masing centroid
- Setiap data memilih centroid yang terdekat
- Tentukan posisi centroid yang baru dengan cara menghitung nilai rata-rata dari data-data yang terletak pada centroid yang sama
- Jika posisi centroid baru dengan centroid yang lama masih belum memiliki kesamaan maka diulang lagi pada langkah ke 3 sampai menemukan kesamaan.
- Tentukan jumlah klaster yang akan dibentuk dan k sebagai nilainya
- Hitung jarak setiap inputan terhadap *centroid – centroid* menggunakan persamaan *Euclidean*, dengan rumus  $i$

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

### 3. HASIL DAN PEMBAHASAN

Data Tindak pidana korupsi di Indonesia ditangani oleh 33 wilayah di tahun 2014 dan semua data tindak pidana Korupsi masing-masing wilayah akan dijadikan contoh pemrosesan dengan cara *Data Mining*. Format data tersebut berupa data Microsoft Excel dan dilakukan proses pengujian menggunakan tool aplikasi *rapidminer* yang bisa mengolah data file tipe .xls. Dari hasil pengolahan data penelitian ini menggunakan data tindak pidana korupsi tahun 2014 yang dapat dilihat dalam tabel berikut.

Tabel ke-1 Data Master Tindak Pidana Korupsi Tahun 2014

No	Wilayah	SPDP	Putusan	Belum Selesai
1	Sumatra1	140	91	49
2	Bali1	6	2	4
3	Sumatra2	4	2	2

No	Wilayah	SPDP	Putusan	Belum Selesai
4	Jawa1	8	8	0
5	Sumatra3	50	37	13
6	Jawa2	18	12	6
7	Jawa3	41	24	17
8	Sulawesi1	7	6	1
9	Sumatra4	97	46	51
10	Jawa4	26	20	6
11	Jawa5	55	33	22
12	Jawa6	41	25	16
13	Kalimantan1	28	14	14
14	Kalimantan2	29	24	5
15	Kalimantan3	11	9	2
16	Kalimantan4	13	4	9
17	Sumatra5	4	4	0
18	Sumatra6	25	17	8
19	Papua1	31	11	20
20	Papua2	17	8	9
21	Bali2	8	5	3
22	Bali3	37	24	13
23	Papua3	4	0	4
24	Papua4	1	1	0
25	Sumatra7	78	55	23
26	Sulawesi2	3	2	1
27	Sulawesi3	12	9	3
28	Sulawesi4	8	5	3
29	Sulawesi5	10	3	7
30	Sulawesi6	15	8	7
31	Sulawesi7	95	56	39
32	Sumatra8	107	69	38
33	Sumatra9	73	49	24

Data yang akan digunakan penelitian ini adalah Jumlah Kasus, Kasus Selesai, untuk menentukan berapa jumlah cluster dan awal titik *cluster* pusat secara acak.

Dari hasil tersebut, bisa disimpulkan dari tiap-tiap titik awal *cluster* yang berdasarkan urutan. Untuk menentukan wilayah yang terjadi tindak pidana korupsi tidak rawan pidana korupsi sedang dan rawan tindak pidana korupsi sebagai berikut:

- Cluster* pusat I: Data ke-1 Wilayah Sumater 1 dengan SPDP = 140 dan Putusan 91
- Cluster* pusat II: Data ke-14 Wilayah Kalimantan2 dengan SPDP = 29 dan Putusan 24
- Cluster* pusat III: Data ke-29 Wilayah Sulawesi5 dengan SPDP = 10 dan Putusan 3.

Untuk hasil tersebut diilustrasikan sebagai berikut:<sup>2</sup>

- Iterasi I
  - Dt-1 ke Cluster Pusat I
 
$$= \sqrt{(140 - 140)^2 + (91 - 91)^2}$$

$$= 0$$
  - Dt-2 ke Cluster Pusat I
 
$$= \sqrt{(6 - 140)^2 + (2 - 91)^2}$$

$$= 160,863296$$
  - Dt-3 ke Cluster Pusat I

$$\begin{aligned}
 &= \sqrt{(4 - 140)^2 + (2 - 91)^2} \\
 &= 162,5330736 \\
 \text{Dt-4 ke Cluster Pusat I} \\
 &= \sqrt{(8 - 140)^2 + (8 - 91)^2} \\
 &= 155,9262646 \\
 \text{Dt-5 ke Cluster Pusat I} \\
 &= \sqrt{(50 - 140)^2 + (37 - 91)^2} \\
 &= 104,9571341 \\
 \text{Dt-6 ke Cluster Pusat I} \\
 &= \sqrt{(18 - 140)^2 + (12 - 91)^2} \\
 &= 145,3444185 \\
 \text{Dt-7 ke Cluster Pusat I} \\
 &= \sqrt{(41 - 140)^2 + (24 - 91)^2} \\
 &= 119,540788 \\
 \text{Dt-8 ke Cluster Pusat I} \\
 &= \sqrt{(7 - 140)^2 + (6 - 91)^2} \\
 &= 157,8416928 \\
 \text{Dt-9 ke Cluster Pusat I} \\
 &= \sqrt{(97 - 140)^2 + (46 - 91)^2} \\
 &= 62,24146528 \\
 \text{Dt-10 ke Cluster Pusat I} \\
 &= \sqrt{(26 - 140)^2 + (20 - 91)^2} \\
 &= 134,3018987 \\
 \text{Dt-11 ke Cluster Pusat I} \\
 &= \sqrt{(55 - 140)^2 + (33 - 91)^2} \\
 &= 102,9028668 \\
 \text{Dt-12 ke Cluster Pusat I} \\
 &= \sqrt{(41 - 140)^2 + (25 - 91)^2} \\
 &= 118,9831921 \\
 \text{Dt-13 ke Cluster Pusat I} \\
 &= \sqrt{(28 - 140)^2 + (14 - 91)^2} \\
 &= 135,915414 \\
 \text{Dt-14 ke Cluster Pusat I} \\
 &= \sqrt{(29 - 140)^2 + (24 - 91)^2} \\
 &= 129,6533841 \\
 \text{Dt-15 ke Cluster Pusat I} \\
 &= \sqrt{(11 - 140)^2 + (9 - 91)^2} \\
 &= 152,8561415 \\
 \text{Dt-16 ke Cluster Pusat I} \\
 &= \sqrt{(13 - 140)^2 + (4 - 91)^2} \\
 &= 153,9415473 \\
 \text{Dt-17 ke Cluster Pusat I} \\
 &= \sqrt{(4 - 140)^2 + (4 - 91)^2} \\
 &= 161,4465856 \\
 \text{Dt-18 ke Cluster Pusat I} \\
 &= \sqrt{(25 - 140)^2 + (17 - 91)^2} \\
 &= 136,7515996 \\
 \text{Dt-19 ke Cluster Pusat I} \\
 &= \sqrt{(31 - 140)^2 + (11 - 91)^2} \\
 &= 135,2072483 \\
 \text{Dt-20 ke Cluster Pusat I} \\
 &= \sqrt{(17 - 140)^2 + (8 - 91)^2} \\
 &= 148,3846353 \\
 \text{Dt-21 ke Cluster Pusat I} \\
 &= \sqrt{(8 - 140)^2 + (5 - 91)^2}
 \end{aligned}$$

$$\begin{aligned}
 &= 157,5436447 \\
 \text{Dt-22 ke Cluster Pusat I} \\
 &= \sqrt{(37 - 140)^2 + (24 - 91)^2} \\
 &= 122,8739191 \\
 \text{Dt-23 ke Cluster Pusat I} \\
 &= \sqrt{(4 - 140)^2 + (0 - 91)^2} \\
 &= 163,6367929 \\
 \text{Dt-24 ke Cluster Pusat I} \\
 &= \sqrt{(1 - 140)^2 + (1 - 91)^2} \\
 &= 165,5928742 \\
 \text{Dt-25 ke Cluster Pusat I} \\
 &= \sqrt{(78 - 140)^2 + (55 - 91)^2} \\
 &= 71,69379332 \\
 \text{Dt-26 ke Cluster Pusat I} \\
 &= \sqrt{(3 - 140)^2 + (2 - 91)^2} \\
 &= 163,370744 \\
 \text{Dt-27 ke Cluster Pusat I} \\
 &= \sqrt{(12 - 140)^2 + (9 - 91)^2} \\
 &= 152,0131573 \\
 \text{Dt-28 ke Cluster Pusat I} \\
 &= \sqrt{(8 - 140)^2 + (5 - 91)^2} \\
 &= 157,5436447 \\
 \text{Dt-29 ke Cluster Pusat I} \\
 &= \sqrt{(10 - 140)^2 + (3 - 91)^2} \\
 &= 156,9840756 \\
 \text{Dt-30 ke Cluster Pusat I} \\
 &= \sqrt{(15 - 140)^2 + (8 - 91)^2} \\
 &= 150,0466594 \\
 \text{Dt-31 ke Cluster Pusat I} \\
 &= \sqrt{(95 - 140)^2 + (8 - 91)^2} \\
 &= 57,00877125 \\
 \text{Dt-32 ke Cluster Pusat I} \\
 &= \sqrt{(107 - 140)^2 + (69 - 91)^2} \\
 &= 39,66106403 \\
 \text{Dt-33 ke Cluster Pusat I} \\
 &= \sqrt{(73 - 140)^2 + (49 - 91)^2} \\
 &= 79,07591289
 \end{aligned}$$

Dari perhitungan tersebut diatas yang menggunakan persamaan *Euclidian Distance* terhadap iterasi I tersebut diatas dapat diamati di tabel ini:

Tabel ke-2 Iterasi I				
No	Wilayah	C1	C2	C3
1	Sumatra1	0,00	129,65	156,98
2	Bali1	160,86	31,83	4,12
3	Sumatra2	162,53	33,30	6,08
4	Jawa1	155,93	26,40	5,39
5	Sumatra3	104,96	24,70	52,50
6	Jawa2	145,34	16,28	12,04
7	Jawa3	119,54	12,00	37,44
8	Sulawesi1	157,84	28,43	4,24
9	Sumatra4	62,24	71,47	97,05
10	Jawa4	134,30	5,00	23,35
11	Jawa5	102,90	27,51	54,08
12	Jawa6	118,98	12,04	38,01
13	Kalimantan1	135,92	10,05	21,10
14	Kalimantan2	129,65	0,00	28,32

No	Wilayah	C1	C2	C3
15	Kalimantan3	152,86	23,43	6,08
16	Kalimantan4	153,94	25,61	3,16
17	Sumatra5	161,45	32,02	6,08
18	Sumatra6	136,75	8,06	20,52
19	Papua1	135,21	13,15	22,47
20	Papua2	148,38	20,00	8,60
21	Bali2	157,54	28,32	2,83
22	Bali3	122,87	8,00	34,21
23	Papua3	163,64	34,66	6,71
24	Papua4	165,59	36,24	9,22
25	Sumatra7	71,69	57,98	85,60
26	Sulawesi2	163,37	34,06	7,07
27	Sulawesi3	152,01	22,67	6,32
28	Sulawesi4	157,54	28,32	2,83
29	Sulawesi5	156,98	28,32	0,00
30	Sulawesi6	150,05	21,26	7,07
31	Sulawesi7	57,01	73,35	100,17
32	Sumatra8	39,66	90,05	117,32
33	Sumatra9	79,08	50,61	78,01

Dan untuk tahap selanjutnya ditentukan *cluster* pusat yang jaraknya terpendek terhadap awal *cluster* yang pemilihannya dilakukan secara acak, dimana untuk jaraknya terpendek dengan *cluster* ditentukan dengan rumus  $MIN(C_{ij} : C_{kj} : C_{lj})$  yang mana untuk menentukan hasil jarak terpendek tersebut terhadap *cluster* juga bisa dilihat sesuai tabel dibawah

Tabel ke-3 Biner jarak terdekat data terhadap Cluster

Data Ke	Wilayah	C1	C2	C3	Cluster
1	Sumatra1	1	0	0	C1
2	Bali1	0	0	1	C3
3	Sumatra2	0	0	1	C3
4	Jawa1	0	0	1	C3
6	Sumatra3	0	0	1	C3
7	Jawa2	0	1	0	C2
8	Jawa3	0	0	1	C3
9	Sulawesi1	1	0	0	C1
10	Sumatra4	0	1	0	C2
11	Jawa4	0	1	0	C2
12	Jawa5	0	1	0	C2
13	Jawa6	0	1	0	C2
14	Kalimantan1	0	1	0	C2
15	Kalimantan2	0	0	1	C3
16	Kalimantan3	0	0	1	C3
17	Kalimantan4	0	0	1	C3
18	Sumatra5	0	1	0	C2
19	Sumatra6	0	1	0	C2
20	Papua1	0	0	1	C3
21	Papua2	0	0	1	C3
22	Bali2	0	1	0	C2
23	Bali3	0	0	1	C3
24	Papua3	0	0	1	C3
25	Papua4	0	1	0	C2
26	Sumatra7	0	0	1	C3
27	Sulawesi2	0	0	1	C3
28	Sulawesi3	0	0	1	C3
29	Sulawesi4	0	0	1	C3
30	Sulawesi5	0	0	1	C3
31	Sulawesi6	1	0	0	C1

32	Sulawesi7	1	0	0	C1
33	Sumatra8	0	1	0	C2

Berdasarkan hasil minimum nilai yang didapat pada penggunaan rumus dan hasilnya juga dapat dilihat maka dapat dikelompokan seperti berikut:

Tabel ke-4 Hasil iterasi / pengelompokan

Cluster Group	Data Cluster	Jumlah Data
1	Wilayah 1,9,31,32	4
2	Wulayah 5,7,10,11,12,13,14,18,19,22,25,33	12
3	Wilayah 2,3,4,6,8,15,16,17,20,21,23,24,26,27,28,30	17

Selanjutnya untuk iterasi ke-2 s/d iterasi terakhir bisa digambarkan sesuai tabel berikut:

Tabel 5 Iterasi biner ke 4

Data Ke	Wilayah	Jarak Centroid		
		C1	C2	C3
1	Sumatra1	0	0	1
2	Bali1	0	1	0
3	Sumatra2	0	0	1
4	Jawa1	0	1	0
5	Sumatra3	0	0	1
6	Jawa2	0	1	0
7	Jawa3	0	0	1
8	Sulawesi1	0	1	0
9	Sumatra4	1	0	0
10	Jawa4	0	1	0
11	Jawa5	0	0	1
12	Jawa6	0	0	1
13	Kalimantan1	0	0	1
14	Kalimantan2	0	1	0
15	Kalimantan3	0	1	0
16	Kalimantan4	0	1	0
17	Sumatra5	0	0	1
18	Sumatra6	0	1	0
19	Papua1	0	1	0
20	Papua2	0	1	0
21	Bali2	0	1	0
22	Bali3	0	1	0
23	Papua3	0	1	0
24	Papua4	0	0	1
25	Sumatra7	0	0	1
26	Sulawesi2	0	0	1
27	Sulawesi3	0	1	0
28	Sulawesi4	0	1	0
29	Sulawesi5	0	1	0
30	Sulawesi6	0	1	0
31	Sulawesi7	1	0	0
32	Sumatra8	1	0	0
33	Sumatra9	0	0	1

Hasil evaluasi berupa Davies Bouldin Index (DBI) sudah dituliskan pada bagian ini. 2 kluster 0,316 sementara pada 3 kluster 0,113 dan pemodelan 4 kluster 0,618. Dengan demikian model terbaik adalah pada terbentuknya 3 kluster dengan nilai DBI sebesar

0,113. DBI terkecil menyatakan bahwa model memiliki bentuk klaster terbaik sebanyak 3.. Untuk membuktikan hasil dari pengelompokan bisa dilihat di tabel yang jarak / C3 adalah wilayah rawan terjadinya korupsi berada di wilayah dengan kategori rawan tindak pidana korupsi berada pada pulau sumatra ada 5 wilayah, pulau jawa ada 3 wilayah dan pulau sulawesi dan kalimantan ada 1 wilayah.

#### 4. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan tersebut diatas, dapat disimpulkan diantaranya:

1. Pengelompokan daerah rawan tindak pidana korupsi di wilayah indonesia belum bisa dijadikan, dikarenakan belum ada inovasi untuk memetakan wilayah yang dianggap sebagai kelompok rawan dan saat ini pengelompokan belum secara modern.
2. Pembagian kategori yang didapat dari data-data SPDP, penuntutan kemudian kategori tersebut dirubah menjadi wilayah tidak rawan tindak pidana korupsi, cukup dan rawan sebagai nilai yang bisa dihasilkan pada masing-masing cluster.

#### 5. DAFTAR PUSTAKA

- [1] R. Rini and L. Damiati, 2017. "Analisis Hasil Audit Pemerintahan dan Tingkat Korupsi Pemerintahan Provinsi di Indonesia," *J. Din. Akunt. dan Bisnis*, vol. 4, no. 1, pp. 73–90, doi: 10.24815/jdab.v4i1.4933.
- [2] R. Nooraeni, H. D. Sariyanti, A. F. F. Iskandar, S. F. Munawwaroh, S. Pertiwi, and Y. Ronaldias, 2020. "Analisis Sentimen Data Twitter Mengenai Isu RUU KPK Dengan Metode Support Vector Machine (SVM)," *Paradig. - J. Komput. dan Inform.*, vol. 22, no. 1, pp. 55–60, 2020, doi: 10.31294/p.v22i1.6869.
- [3] Hernawati and W. Gata, 2019. "Sentimen Analisis Operasi Tangkap Tangan KPK Menurut Masyarakat Menggunakan Algoritma Support Vector Machine, Naive Bayes Berbasis Particle Swarm Optimizition," *Fakt. Exacta*, vol. 12, no. 3, pp. 230–243, doi: 10.30998/faktorexacta.v12i3.4992.
- [4] R. Hablum, A. Khairan, and R. Rosihan, 2019. "Clustering Hasil Tangkap Ikan Di Pelabuhan Perikanan Nusantara (Ppn) Ternate Menggunakan Algoritma K-Means," *JIKO (Jurnal Inform. dan Komputer)*, vol. 2, no. 1, pp. 26–33, doi: 10.33387/jiko.v2i1.1053.
- [5] J. J. Purnama, 2019. "Analisa Algoritma K-Means Clustering Pemetaan Jumlah Tindak Pidana," *Klik - Kumpul. J. Ilmu Komput.*, vol. 6, no. 2, p. 128, doi: 10.20527/klik.v6i2.208.
- [6] L. Suriani, 2020. "Pengelompokan Data Kriminal Pada Poldasu Menentukan Pola Daerah Rawan Tindak Kriminal Menggunakan Data Mining Algoritma K-Means Clustering," *J. Sist. Komput. dan Inform.*, vol. 1, no. 2, p. 151, doi: 10.30865/json.v1i2.1955.
- [7] U. T. Suryadi and Y. Supriatna, 2019. "Sistem Clustering Tindak Kejahatan Pencurian Di Wilayah Jawa Barat Menggunakan Algoritma K-Means," *J. Teknol. Inf. dan Komun.*, no. April, pp. 1–13.
- [8] W. Astuti, A. Widodo, J. T. Elektro, F. Teknik, and U. N. Semarang, 2016. "Pemetaan Tindak Kejahatan Jalanan di Kota Semarang Menggunakan Algoritma K-Means Clustering," vol. 8, no. 1, pp. 5–7.
- [9] I. T. Utami, R. Rais, and W. Seftiani, 2019. "Analisis Klaster Pautan Lengkap Untuk Mengelompokkan Kabupaten/Kota Di Provinsi Sulawesi Tengah Berdasarkan Indikator Kriminalitas," *J. Ilm. Mat. Dan Terap.*, vol. 16, no. 1, pp. 79–88, doi: 10.22487/2540766x.2019.v16.i1.12757.
- [10] M. S. Fajar, 2015. "DI KOTA SEMARANG MENGGUNAKAN METODE K-MEANS CLUSTERING,".
- [11] I. Method, K. C. Based, S. Value, W. Interface, C. Study, and I. U. M. M. Magelang, 2015. "Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface di UMM" *Jurusan Teknik Informatika* vol. 18, no. 1, pp. 76–82.