

## IMPLEMENTASI METODE K-MEANS DALAM PENGKLASTERAN DAERAH PUNGUTAN LIAR DI KABUPATEN SUKABUMI (STUDI KASUS : DINAS KEPENDUDUKAN DAN PENCATATAN SIPIL)

Falentino Sembiring<sup>1</sup>, Octaviana<sup>2</sup>, Sudin Saepudin<sup>3</sup>

<sup>1,2,3</sup>Universitas Nusa Putra

<sup>1</sup>Falentino.sembiring@nusaputra.ac.id

<sup>2</sup>Octaviana366@gmail.com

<sup>3</sup>Sudin@nusaputra.ac.id

**Abstrak** - Pungutan liar (pungli) merupakan tindakan kejahatan yang dilakukan oleh pihak yang tidak bertanggung jawab atau seseorang atau pegawai negeri atau pejabat negara dengan cara meminta pembayaran uang yang tidak termasuk kedalam aturan administrasi yang dibutuhkan. Metode K-Means dapat membantu mengklasifikasikan daerah pungutan liar di kabupaten Sukabumi pada dinas kependudukan dan pencatatan sipil. Variabel yang digunakan dalam penelitian ini E-KTP, AKTA dan Kartu Keluarga Setiap variabel memiliki nilai bobot yang berbeda. Data yang akan digunakan dalam penelitian ini adalah data laporan pada bulan Januari 2019 berbentuk SQL. Penelitian ini hanya akan membahas tentang berapa banyak kejadian pungli yang terjadi di setiap kecamatan yang ada di kabupaten Sukabumi tentang kependudukan dan pencatatan sipil. Adapun tujuan dari penelitian ini adalah untuk menentukan *cluster* tingkat pungutan liar tinggi, sedang dan rendah. Hasil dari penelitian ini memperoleh data indeks kecamatan dengan tingkat jumlah laporan masyarakat daerah terhadap pungutan liar, data dengan klasifikasi tingkat tinggi yaitu Cireunghas, Gegerbitung, Kalapa Nunggal, Kalibunder, Purabaya, Simpenan, Parung Kuda, Sukaraja, Nagrak, Nyalindung, Pelabuhanratu, Surade, Warungkiara. Data tingkat pungutan liar sedang terdapat pada 19 kecamatan dan 15 tingkat pungutan liar rendah. Hal ini dapat menjadi masukan kepada dinas kependudukan dan pencatatan sipil, kecamatan yang menjadi tindakan pungutan liar paling tinggi diprioritas untuk melakukan penyuluhan kepada kecamatan tersebut.

**Kata kunci** : Pungutan Liar, K-Means, Data Mining, Clustering

**Abstract** - *Illegal Levies is a criminal offense done by someone or civil servant and the government by demanding a specific amount of money outside the right protocols or policy that does not related with the official administration fees. K-Means method will be able to help for classifying the illegal levies area in sukabumi district at The Department of Population and Civil Registration. Variables that used in this research are electronic ID card, birth certificate and family card. Meanwhile, data that will be used in this research are report data on January 2016. The purpose of this research is to determine the rate of illegal levies. The result that have been collected based on the high extortion rate from each sub-districts are Cireunghas, Gegerbitung, Kalapa Nunggal, Kalibunder, Purabaya, Simpenan, Parung Kuda, Sukaraja, Nagrak, Nyalindung, Pelabuhanratu, Surade, Warungkiara, 19 sub-districts with medium rate of illegal levies and 15 sub-district with low rate of illegal levies. This can become a recommendation for The Department of Population and Civil Registration, sub-districts with the highest rate of illegal levies will be prioritized in counseling for the certain sub-districts.*

**Keywords** : *Illegal levies , K-Means, Data Mining, Clustering*

### 1. PENDAHULUAN

Indonesia memiliki jumlah penduduk terpadat ke 4 di dunia dengan jumlah penduduk 260.580.739 juta jiwa. Indonesia menduduki posisi ke 4 setelah urutan pertama diduduki oleh negara Tiongkok. Menurut UU RI tentang Administrasi Kependudukan Nomor 23 Tahun 2006 menyatakan setiap warga Indonesia hanya diperbolehkan 1 Kartu Tanda Penduduk pada Nomor Induk Kependudukan (NIK) NIK merupakan identitas tunggal setiap penduduk dan berlaku seumur hidup untuk warga negara Indonesia dan untuk warga asing disesuaikan dengan masa berlaku izin tinggal tetap (*Seberapa Pentingkah NIK, Data Kependudukan dan KTP Elektronik?*, n.d.)

Salah satu permasalahan yang muncul adalah di kabupaten Sukabumi masih terjadi pungutan liar pada administrasi kependudukan. Di setiap kecamatan terdapat banyak calo sebagai perantara untuk mengurus administrasi kependudukan dengan harga sesuai tarif dan kesepakatan.

Menurut Ombudsman (2017) bentuk pungutan liar yang paling banyak terjadi di calo pengurusan sebesar 52,17%, Calo antrean 19,57%, permintaan biaya pembuatan 15,22%, dan 2,17% kotak sumbangan seikhlasnya. (*Calo dan Pungli Masih Marak, Ombudsman RI: Masyarakat Tak Mau Repot dan Pilih Jalan Pintas - Ombudsman RI*, n.d.) Dengan permasalahan tersebut peneliti ingin mengelompokkan daerah pungutan liar di kabupaten

Implementasi Metode K-Means dalam Pengklasteran Daerah Pungutan Liar di Kabupaten Sukabumi  
(Studi Kasus : Dinas Kependudukan dan Pencatatan Sipil)

Sukabumi dengan menggunakan metode K-Means sebagai perhitungan penentuan 3 klaster yaitu tingkat tinggi, sedang, dan rendah.

Pada tahun (2016) Kumala dkk melakukan penelitian tentang “Perbandingan K-Means Dan Fuzzy C-Means Clustering Pada Metode Radian Basis Function Neural Network (RBFNN) Untuk Klasifikasi Stadium Kanker Payudara”. Hasil dari penelitian tersebut mereka menyimpulkan bahwa metode K-Means clustering menggunakan 3 parameter data untuk menentukan tingkat nilai sensitivitas 75%, spesifisitas 93,75%, dan akurasi 80,21% . (Kumala & Wutsqa, 2016)

Dengan adanya penelitian ini, penulis berharap permasalahan tentang pungutan liar bisa dapat dinas kependudukan dan pencatatan sipil evaluasi kembali dengan mengadakan program kerja penyuluhan pada tiap kecamatan yang masuk ke dalam klaster daerah pungutan liar tinggi.

**A. Identifikasi Masalah**

Peneliti melihat bahwa pungutan liar di kabupaten Sukabumi masih sering terjadi dan hal tersebut membuat warga masyarakat merasa dirugikan secara finansial karena adanya pembayaran yang tidak dimintai oleh pihak terkait dan tidak tercantum dalam prosedur.

**B. Rumusan Masalah**

Dalam penelitian masalah yang dirumuskan adalah bagaimana menerapkan metode K-Means untuk mengklaster daerah pungutan liar di kabupaten Sukabumi.

**C. Batasan Masalah**

Penelitian ini berfokus bagaimana tingkatan (tingkat tinggi, sedang, dan rendah) terhadap pungutan liar di kabupaten Sukabumi pada kependudukan dan pencatatan sipil.

**D. Tujuan**

Untuk membangun sistem pengelompokan daerah pungutan liar ke dalam tiga klaster di kabupaten Sukabumi pada bidang kependudukan dan pencatatan sipil.

**E. Manfaat**

Semua Sebagai bahan evaluasi untuk melakukan program kerja yang telah dijalankan oleh pihak dinas kependudukan dan pencatatan sipil dan memberikan kemudahan untuk mengetahui daerah yang masih terjadinya pungutan liar di kabupaten Sukabumi.

**2. TINJAUAN PUSTAKA**

**A. Penelitian Terkait**

Agus Perdana Windarto (2017) melakukan penelitian pengeksporan buah yang berstudi kasus di Direktorat Jenderal Bea dan Cukai dengan judul “Penerapan Data Mining Menggunakan K-Means Clustering” dari penelitian tersebut agus

menyimpulkan bahwa tingkat **cluster tinggi** yakni dua negara , dua negara **cluster sedang** dan enam negara **cluster tingkat rendah** (Windarto, 2017)

Tabel 1. Nilai Bobot Setiap Variabel

<b>Tinggi</b>	1. India 2. Pakistan
<b>Sedang</b>	1. Singapura 2. Bangladesh
<b>Rendah</b>	1. hongkong 2. Tiongkok 3. Iran 4. Malaysia 5. Nepal 6. Vietnam

Arsya Yoga Pratama, Iman Ardhi Prabowo, dan Alifah Normassari melakukan penelitian tentang permasalahan sering terjadi pungutan liar di Kabupaten Kudus. Penelitian tersebut diteliti pada tahun 2017 tanpa menggunakan metode clustering k-means. Mereka melakukan pembuatan sistem berbasis android untuk melaporkan kejadian pungutan liar. Kesimpulan yang penulis dapat pada aplikasi yang mereka buat terdapat kelebihan yaitu masyarakat nantinya dapat memberikan bukti kegiatan pungutan liar berupa audio, gambar ataupun video serta keakurasian tempat terjadinya pungutan liar (Pratama et al., 2017)

**B. Data Mining**

*Data mining* adalah proses pencarian pola yang penting dari data yang ada didalam *database* yang besar (Rante Rerung, 2018)

*Data mining* merupakan istilah *pattern recognition* merupakan algoritma untuk pengolahan data guna menemukan pola data menjadi pengetahuan baru . Data yang diolah dengan teknik data *mining* akan menghasilkan suatu pengetahuan bersumber dari data lama, hasilnya untuk menentukan keputusan bisnis. (Sembiring & Sari, 2019)

*Data mining* dibagi menjadi 2 kategori utama, yaitu:

1. *Descriptive mining*, tahapan untuk menemukan karakteristik penting dari data dalam satu basis data.
2. *Predictive*, tahapan untuk mempelajari bagaimana *pattern design* dari data set yang berupa variabel.

Teknik pengolahan data dengan bantuan algoritma *data mining*. pengolahan dilakukan dengan membangun sebuah *pattern design*, kemudian model tersebut membentuk pola data yang lain yang tidak berada dalam basis data . Tujuannya untuk memperoleh *pattern design* untuk menjadikan data sebagai grafik untuk menjadi pengetahuan. (Gustian et al., 2018)

**C. Metode K-Means**

K-means merupakan metode pengklasteran secara *partitioning* yang memisahkan data ke dalam kelompok yang berbeda. Dengan *partitioning* secara iteratif, K-Means mampu meminimalkan rata-rata jarak setiap data ke klasternya. Dalam algoritma K-Means, setiap data harus termasuk ke *cluster* tertentu pada suatu tahapan proses, pada tahapan proses berikutnya dapat berpindah ke *cluster* yang lain

$$v_{ij} = \frac{1}{N_k} \sum_{i=1}^{N_k} X_{ij}$$

Rumus Penentuan *Centroid* Baru

K-Means merupakan algoritma yang termasuk ke dalam data *mining* bidang *Clustering* atau pengklasteran data. K-Means *Clustering* adalah suatu metode penganalisaan data yang melakukan proses pemodelan tanpa supervisi dan merupakan metode yang termasuk mengelompokkan data secara sistem partisi.

K-Means *Clustering* adalah algoritma yang termasuk dalam klasifikasi terhadap objek berdasarkan atribut. Pemartisian ini digunakan untuk mencari nilai jarak minimum antara data dan nilai *centroid* yang telah di set baik secara *random* dengan cara menentukan nilai *centroid* berdasarkan K yang berupa *object* yang berurutan. *Centroid* merupakan hasil rata-rata dari sebuah bentuk objek dari seluruh titik dalam objek.

**D. Tahapan metode K-Means**

Di dalam metode K-Means memiliki beberapa tahapan yang harus dilakukan saat melakukan perhitungan untuk mencari data *cluster* hingga iterasi menjadi sama dengan hitungan yang sebelumnya, berikut adalah tahapan K-Means :

- a. Menentukan **k** sebagai jumlah *cluster* (C1, C2 dan C3) yang ingin dibentuk.
- b. Tentukan nilai untuk pusat *cluster* (*centroid*).
- c. Hitung jarak setiap data *input* terhadap masing-masing *centroid* menggunakan rumus *Euclidean Distance* hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*. Di bawah ini adalah persamaan *Euclidean Distance* :

$$d(x | j, y_j) = \sqrt{\sum_{i=1}^n (x_j - y_j)^2}$$

**Keterangan :**

- d(x,y) : jarak antara data pada posisi titik x dan y
- x :Posisi titik data pertama (pusat klaster)
- y :posisi titik data kedua (data dari N)
- n : jumlah atribut data

Jarak terpendek antara *centroid* dengan *object* menentukan titik klaster antar *object*. Adapun rumus iterasi lainnya didefinisikan sebagai berikut :

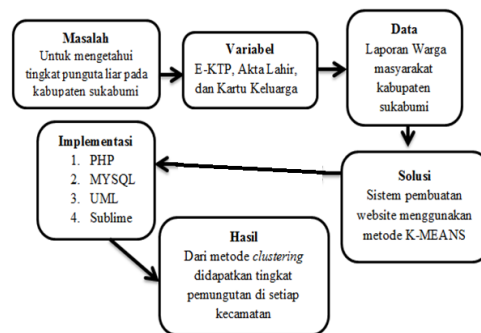
**Keterangan :**

- $v_{ij}$  = *Centroid*, *cluster* pertama pada variabel ke-j
- $N_k$  = Jumlah data yang menjadi anggota *cluster* ke-i
- i,k = Indeks dari *cluster*
- j = Indeks dari variabel
- $X_{ij}$  = Nilai data ke-k yang ada di dalam *cluster* ini untuk variable ke-j

- d. Kemudian lakukan iterasi selama partisi data masih bergerak (tidak ada lagi objek yang bergerak ke partisi lain), bila masih maka ke poin 3.
- e. Bila grup data sekarang sama dengan grup data sebelumnya, maka hentikan iterasi. Data telah dipartisi sesuai nilai *centroid* akhir.

**E. Kerangka Berpikir**

Adapun kerangka berpikir dalam penelitian ini adalah sebagai berikut ini :



Gambar 1. Kerangka berpikir

**3. METODE PENELITIAN**

**A. Tahap Pengolahan Data**

Pengolahan data dilakukan untuk mengetahui hasil dari data yang diteliti. Data yang telah ada diproses menggunakan metode K-Means untuk mengklaster data pemungutan liar pada setiap kecamatan dan Kabupaten di Sukabumi.

Sebelum ke tahap perhitungan K-Means data SQL akan dipilih untuk dilakukan *filtering*, setelah data *cleaning* maka dilakukan pemasukan data nilai bobot untuk proses perhitungan ke tahap selanjutnya. Di bawah ini tabel nilai bobot yang diberikan pihak dinas kependudukan untuk setiap kriteria:

Implementasi Metode K-Means dalam Pengklasteran Daerah Pungutan Liar di Kabupaten Sukabumi  
(Studi Kasus : Dinas Kependudukan dan Pencatatan Sipil)

Tabel 2. Nilai Bobot Setiap Variabel

Variabel	Nilai Bobot
Kartu Keluarga	3
Akta	2
E-KTP	1

**B. Tahapan Clustering**

Dalam menentukan *cluster* dibutuhkan sebuah *flowchart* untuk menentukan alur perhitungan. Untuk lebih jelasnya dapat dilihat *flowchart* berikut ini



Gambar 2. Flowchart Tahapan pengelompokan data

**C. Tahapan Analisis data**

Penulis melakukan analisis pada data pemungutan liar di Kabupaten Sukabumi untuk upaya pengelompokan atau mengklaster data pemungutan liar di setiap kecamatan kabupaten Sukabumi. Data yang sudah diperoleh dari data laporan warga yang berbentuk data SQL dihitung dengan nilai bobot selanjutnya diolah menggunakan perhitungan K-Means (Ghosh & Kumar, 2013) Melakukan *clustering*, data yang didapat akan dihitung terlebih dahulu berdasarkan banyaknya kecamatan untuk upaya melihat pemungutan liar yang paling banyak dilakukan. Kriteria yang digunakan yaitu sebanyak 3 kriteria penilaian yaitu Kartu keluarga(3), Akta (2) dan E-KTP(1) dapat dilihat pada tabel 2 data pemungutan liar.

Tabel 3. Data Pungutan Liar

No	Kecamatan	E-KTP	Akta Kelahiran	Kartu Keluarga
1	Bojong Genteng	1	5	4
2	Cidolog	1	1	3
3	Cireunghas	5	2	6
4	Kebonpedes	2	1	2
5	Lengkong	3	4	1
....	.....	.....	.....	.....
47	Cisolok	2	6	5

**D. Centroid Data**

*Centroid* merupakan titik pusat data untuk menghitung rata-rata vector sebagai centroid. Dalam penerapan algoritma K-means dihasilkan nilai titik tengah atau centroid dari data yang didapat dengan ketentuan bahwa klasterisasi yang diinginkan adalah 3, Penentuan *cluster* dibagi menjadi 3 bagian yaitu *cluster* tingkat pemungutan liar tinggi (C1), tingkat pemungutan liar sedang (C2), dan tingkat pemungutan liar rendah (C3). Penentuan titik *cluster* (C1) diambil dari data nilai maksimal, untuk *cluster* (C2) nilai data rata-rata, dan *cluster* (C3) diambil dari nilai data minimal atau nilai terendah. Sehingga posisi *centroid* juga terdapat 3 titik. Penentuan nilai titik *cluster* tersebut dapat dilihat Tabel 3 dibawah ini:

Tabel 4. Nilai Centroid Awal

	A	B	C
C1	7	18	24
C2	3	7,468085106	10,53191489
C3	1	2	3

Keterangan :

- A = E-KTP
- B = Akta Kelahiran
- C = Kartu Keluarga

**E. Perhitungan Metode K-Means**

Setelah data sudah ada maka dimasukkan kedalam rumus untuk tahap pengklasteran data, dan dibawah ini adalah contoh perhitungan dari data yang ada setelah dimasukkan ke dalam rumus K-Means:

$$d(x | j, y_j) = \sqrt{\sum_{i=1}^n (x_j - y_j)^2}$$

Sehingga menentukan data *cluster*

$$d(x_1, c_1) = \sqrt{(a_1 - c_{1a})^2 + (b_1 - c_{1b})^2}$$

$$d(x_2, c_1) = \sqrt{(a_2 - c_{1a})^2 + (b_2 - c_{1b})^2}$$

$$d(x_3, c_1) = \sqrt{(a_3 - c_{1a})^2 + (b_3 - c_{1b})^2}$$

$$d(x_4, c_1) = \sqrt{(a_4 - c_{1a})^2 + (b_4 - c_{1b})^2}$$

Sehingga

$$C1 = \sqrt{36 + 64 + 144}$$

$$C1 = \sqrt{244}$$

$$C1 = 15,62049935$$

Dari hasil perhitungan diatas didapatkan hasil bahwa jarak data pertama dengan pusat *cluster* pertama adalah 15,62049935. Pada tahapan selanjutnya menentukan C2 menggunakan rumus seperti berikut ini

$$d(x_1, c_2) = \sqrt{(a_1 - c_{2a})^2 + (b_1 - c_{2b})^2}$$

$$d(x_2, c_2) = \sqrt{(a_2 - c_{2a})^2 + (b_2 - c_{2b})^2}$$

$$d(x_3, c_2) = \sqrt{(a_3 - c_{2a})^2 + (b_3 - c_{2b})^2}$$

$$d(x_4, c_2) = \sqrt{(a_4 - c_{2a})^2 + (b_4 - c_{2b})^2}$$

Sehingga

$$C_2 = \sqrt{4 + 6,41059303045903 + 2,15527389020371}$$

$$= \sqrt{12,5658669206627}$$

$$= 3,544836655$$

Dari hasil perhitungan diatas di dapatkan hasil bahwa jarak data pertama dengan pusat cluster kedua adalah 3,544836655.

$$C_3 = \sqrt{0 + 64 + 81}$$

$$= \sqrt{145}$$

$$= 12,04159458$$

Dari hasil perhitungan diatas didapatkan hasil bahwa jarak data mahasiswa pertama dengan pusat cluster ketiga adalah 12,04159458. Setelah keseluruhan data dihitung maka didapatkan hasil jarak berikut untuk Iterasi 1:

No	Kecamatan	Jarak Ke Centroid			Terdekat
		C1	C2	C3	
1	Bojong Genteng	15,62049935	3,544836655	12,04159458	3,544836655
2	Cidolog	22,737634	6,020524726	6	6
3	Cireunghas	15,3622915	8,473482723	15,65247584	8,473482723
4	Kebonpedes	24,59674775	7,172043456	3,16227766	3,16227766
5	Lengkong	23,60084744	7,550673839	6,32455532	6,32455532
.....	.....	.....	.....	.....	.....
47	Cisolok	11,91637529	6,442207473	15,65247584	6,442207473

Gambar 3. hasil perhitungan literasi pertama

Dari data pemungutan liar yang sudah ada didapatkan pengelompokan pada iterasi 1 untuk 3 cluster tersebut. Cluster tingkat pemungutan liar tinggi (C1) yaitu Prabaya, cluster tingkat pemungutan liar sedang (C2) yaitu ada 35 kecamatan dan cluster tingkat pemungutan liar rendah (C3) yaitu ada 43 kecamatan lainnya dan cluster tingkat pemungutan liar rendah.

Proses pengelompokan data pada iterasi 1 dan Clustering digambarkan pada gambar berikut:

No	Kecamatan	Jarak Ke Centroid		
		C1	C2	C3
1	Bojong Genteng		X	
2	Cidolog			X
3	Cireunghas		X	
4	Kebonpedes			X
5	Lengkong			X
.....	.....	.....	.....	.....
47	Cisolok		X	

Gambar 4. Clustering pengelompokan data



Gambar 5. Pengelompokan Cluster Pertama

Terus dilakukan perhitungan hingga data objek tidak berubah tempat. Pada penelitian ini data yang dihitung berhenti pada iterasi ke enam. (J. O. Ong, 2013)

#### 4. HASIL DAN PEMBAHASAN

##### A. Hasil Penelitian

Hasil penelitian berisi pengelompokan data ke dalam tiap-tiap klaster yang dibutuhkan. hasil pengelompokan data daerah pungutan liar yang telah dihitung menggunakan metode K-Means. Implementasi di dalam sistem pengelompokan data kecamatan tentang pungutan liar didapat dari analisa dan desain sistem pada BAB sebelumnya. Data hasil penelitian dimasukkan secara komputerisasi dalam bentuk pengelompokan daerah pungutan liar yang dibangun menggunakan bahasa pemrograman PHP dan MySQL sebagai database, kemudian hasil perhitungan dan eksekusi sistem dari metode K-Means.

##### B. Hasil Klasifikasi Pungutan Liar

Dari 479 data yang akan dihitung menggunakan metode K-Means. Diketahui bahwa terdapat 13 kecamatan yang masuk kedalam cluster tingkat pungutan liar tinggi (C1) yaitu Cireunghas, Geger Bitung, Kalapa Nunggal, Kalibunder, Purabaya, Simpanan, Parung Kuda, Sukaraja, Nagrak, Nyalindung, Pelabuhanratu, Surade, dan Warungkiara, 19 tingkat pungutan liar sedang (C2) yaitu Bojong Genteng, Sukalarang, Waluran, Bantargadung, Curugkembar, Ciambar, Gunungguruh, Pabuaran, Cimanggu, Cidahu, Ciracap, Cibadak, Kecamatan, Cikembar, Jampang Kulon, Jampang Tengah, Cikidang, Sagaranten, Cicurug, dan Cisolok, dan 16 kecamatan tingkat pungutan liar rendah (C3) yaitu Cidolog, Kebonpedes, Kecamatan, Lengkong, Cibitung, Cidap, Kabandungan, Parakan Salak, Sukabumi, Cicantayan, Tegalbuleud, Caringin, Ciemas, Cikakak, Kadudampit, dan Cisaat pada hasil perhitungan laporan pada bulan januari 2019.

##### C. Implementasi Sistem

###### 1. Halaman Utama

Halaman Utama merupakan tampilan pertama yang ditampilkan khusus untuk pelapor (warga

# Implementasi Metode K-Means dalam Pengklasteran Daerah Pungutan Liar di Kabupaten Sukabumi (Studi Kasus : Dinas Kependudukan dan Pencatatan Sipil)

masyarakat kabupaten Sukabumi) untuk melaporkan kejadian pungutan liar yang pelapor alami.



Gambar 6. Halaman Utama

## 2. Halaman Laporan

Pelapor atau warga kabupaten Sukabumi dapat melaporkan kejadian pungutan liar dengan mengisi data yang telah tertera di *form* yang telah ditentukan oleh pihak dinas. Data diri pelapor tidak akan dipublikasikan atau disebar luaskan karena data pelapor dilindungi oleh perundang-undangan 1945. Berikut tampilan Halaman Laporan:



Gambar 7. Halaman Laporan

## 3. Halaman Login Admin

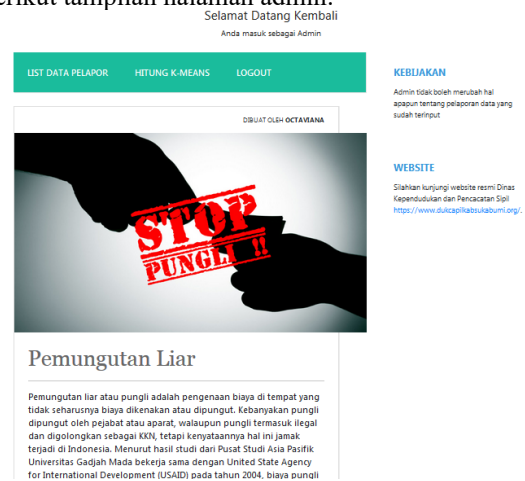
Halaman *Login* merupakan halaman khusus untuk admin agar dapat mengakses data laporan warga yang sudah terinput ke dalam sistem. Akses admin yang diberikan tergantung kesepakatan pihak dinas siapa saja yang dapat mengakses menu halaman *login* admin oleh sebab itu peneliti tidak memberikan registrasi agar tidak terjadi *user* yang bukan pegawai dinas. Berikut tampilan *login* halaman admin:



Gambar 8. Halaman Login Admin

## 4. Halaman Admin

Setelah admin melakukan *login* maka halaman selanjutnya yang ditampilkan adalah halaman admin untuk proses penghitungan data dan melihat jumlah data laporan yang sudah masuk ke dalam sistem. Berikut tampilan halaman admin:



Gambar 9. Halaman Admin

## 5. List Data Pelapor

Halaman ini hanya bisa diakses oleh admin dan admin bersangkutan tidak dapat merubah atau menghapus data dari hasil laporan warga kabupaten Sukabumi atas keluhan pungutan liar yang terjadi atau maraknya calo di kecamatan tersebut, karena apabila di halaman ini admin diberikan akses hapus dan edit maka bisa terjadinya pemalsuan atau penyalpian data. Oleh sebab itu admin hanya dapat melihat sudah berapa banyak laporan yang sudah terinput. Berikut tampilan halaman *list* data laporan warga kabupaten Sukabumi :

No	Nama	Kecamatan	Desa	Tanggal	No Hp	Pelaporan	Komentar
1	UJAL DARUL HAJARI	RELAWAN RAYU	RELAWAN RAYU	05-01-2018	081-888-8101	Kelu keluarga	RIWAYE
2	ADAM LINDA WIDYANINGRUM	RELAWAN RAYU	TONGKONG	05-01-2018	081-821-1189	Kelu keluarga	RIWAYE
3	SALAFIYAH	RELAWAN RAYU	CITRAH	05-01-2018	081-718-4883	laku	RIWAYE
4	BUTIK	RELAWAN RAYU	SUNDRING	04-01-2018	082-626-2058	laku	RIWAYE
5	ROHMAN	RELAWAN RAYU	CEKONG	05-01-2018	083-627-7628	4779	RIWAYE
6	HENDRI LATHAM	RELAWAN RAYU	PANDEUREAN	05-01-2018	082-677-7669	4779	RIWAYE
7	SABRUDIN	RELAWAN RAYU	CIKIDU	07-01-2018	082-771-2762	4779	RIWAYE
8	AL KHARITS	RELAWAN RAYU	CITRAH	08-01-2018	082-881-0201	4779	RIWAYE
9	LIWA	RELAWAN RAYU	CIMANGGU	08-01-2018	083-471-3402	4779	RIWAYE
10	DESI	RELAWAN RAYU	JAWAH	10-01-2018	081-884-9351	4779	RIWAYE

Gambar 10. Halaman List Data Laporan

## 6. Sorting Data

Setelah data laporan warga kabupaten Sukabumi didapat maka data akan di *sorting* sesuai dengan kebutuhan yang ingin dieksekusi selanjutnya. Berikut tampilan halaman *sorting* data :

Gambar 11. Halaman *Sorting* Data

No	Kecamatan	n Ktp	Atas	Kartu Keluarga
1	BANTARAGUNUNG	4	4	3
2	BOJONG GENTENG	4	5	1
3	CAMPUR	3	4	4
4	CIAMBAR	2	4	3
5	CIANDAM	2	4	3
6	CIATUR	2	3	4
7	CIAYURAN	2	3	4
8	CIJURUG	4	5	3
9	CIJURUP	2	2	6
10	CIJAHU	1	5	7
11	CIJOLEG	3	3	1
12	CIKAS	2	3	2
13	CIKAMAH	1	3	5
14	CIKEMBAR	3	5	1
15	CIKANGSI	2	6	7

Gambar 12. Hasil *Sorting* Data

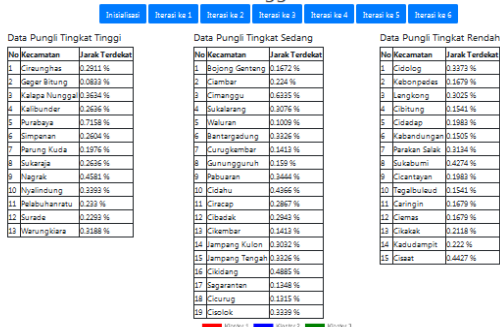
### 7. Nilai Bobot

Apabila data telah ter-*sorting* maka yang selanjutnya dilakukan oleh admin ialah memasukkan nilai bobot untuk setiap variabel yang memiliki nilai bobot yang berbeda sesuai dengan kesepakatan peneliti dan dinas. Berikut tampilan penginputan nilai bobot:

Gambar 13. Nilai Bobot

### 8. Hitung *K-Means*

Klasterisasi Menggunakan *K-means*



Gambar 14. Hasil Perhitungan

Bila nilai bobot sudah dimasukkan maka selanjutnya halaman akan menampilkan hasil perhitungan menggunakan metode *k-means* dengan data yang sudah tersorting sebelumnya. halaman perhitungan menggunakan metode *k-means clustering* untuk penentuan kecamatan mana saja yang masuk ke dalam tiga *cluster* pungutan liar tinggi, sedang, dan rendah ditunjukkan pada Gambar 14.

### D. Pembahasan

Data yang digunakan dalam penelitian ini yaitu hasil laporan warga kabupaten Sukabumi yang masih terjadinya pungutan liar terhadap warga oleh orang tertentu atau melalui calo untuk mengurus pencatatan sipil. Penulis menggunakan metode *k-means* untuk pengelompokan daerah pungutan liar ke dalam tiga golongan.

Dalam penelitian ini, proses analisis yang dilakukan untuk mengelompokkan pungutan liar di setiap kecamatan yang ada di kabupaten Sukabumi untuk periode 1 bulan dan dibantu dengan pihak dinas kependudukan dan pencatatan sipil serta melakukan wawancara. Sistem yang digunakan dalam perhitungan metode *k-means clustering* yaitu php mysql dengan pengkodean menggunakan sublime untuk pengimplementasian sistem dengan menerapkan metode yang peneliti gunakan.

Hasil perhitungan dari sistem tidak jauh berbeda dengan perhitungan manual yaitu sebanyak 13 kecamatan masih marak terjadinya pungutan liar melalui calo setempat di kecamatan tersebut atau para pegawai dinas yang meminta bayaran secara tersirat dan 19 kecamatan masuk ke dalam kelompok tingkat pungutan liar sedang, sementara 15 kecamatan masuk kedalam pungutan liar rendah. Setelah sistem telah selesai dibuat warga dapat melapor tanpa harus melakukan login warga langsung melakukan pelaporan pada sistem dan hasil laporan tersebut akan simpan pada *database*, tahapan selanjutnya akan dilakukan proses pengelompokan menggunakan metode *k-means clustering* oleh admin dinas kependudukan dan pencatatan sipil.

Sistem akan menampilkan grafik pengelompokan data dengan hasil yang sesuai dengan data hasil laporan warga kabupaten Sukabumi terhadap terjadinya pungutan liar. Hasil dari pengelompokan daerah pungutan liar ini dapat digunakan sebagai data evaluasi untuk program kerja penyuluhan terhadap warga kabupaten Sukabumi oleh dinas kependudukan dan pencatatan sipil.

### 5. KESIMPULAN DAN SARAN

Metode *K-Means* dapat diimplementasikan ke dalam sistem dengan menggunakan PHP dan MySQL untuk membantu pengelompokan hasil laporan

Implementasi Metode K-Means dalam Pengklasteran Daerah Pungutan Liar di Kabupaten Sukabumi  
(Studi Kasus : Dinas Kependudukan dan Pencatatan Sipil)

warga terhadap pungutan liar di setiap kecamatan kabupaten Sukabumi.

Hasil penelitian ini memperoleh data indeks kecamatan dengan tingkat jumlah laporan masyarakat daerah terhadap pungutan liar, data dengan kasifikasi tingkat tinggi yaitu Cireunghas, Gegerbitung, Kalapa Nunggal, Kalibunder, Purabaya, Simpenan, Parung Kuda, Sukaraja, Nagrak, Nyalindung, Pelabuhanratu, Surade, Warungkiara. Data tingkat pungutan liar sedang terdapat pada 19 kecamatan dan 15 tingkat pungutan liar rendah. Sebagai saran juga kepada dinas kependudukan dan pencatatan sipil, kecamatan yang menjadi tindakan pungutan liar paling tinggi diprioritaskan untuk melakukan penyuluhan kepada kecamatan tersebut.

## 6. DAFTAR PUSTAKA

- Calo dan Pungli Masih Marak, Ombudsman RI: Masyarakat Tak Mau Repot dan Pilih Jalan Pintas - Ombudsman RI.* (n.d.). Retrieved February 4, 2020, from <https://ombudsman.go.id/perwakilan/news/r/pwk--calo-dan-pungli-masih-marak-ombudsman-ri-masyarakat-tak-mau-repot-dan-pilih-jalan-pintas>
- Ghosh, S., & Kumar, S. (2013). Comparative Analysis of K-Means and Fuzzy C-Means Algorithms. *International Journal of Advanced Computer Science and Applications*, 4(4).  
<https://doi.org/10.14569/ijacsa.2013.040406>
- Gustian, D., Yulitasari, S., Hundayani, D., Muslih, & Nunik. (2018). Comparison of C4.5 method based optimization algorithm to determine eligibility of beneficiaries of direct community assistance (Case study: Kelurahan Cicurug). *3rd International Conference on Computing, Engineering, and Design, ICCED 2017*.  
<https://doi.org/10.1109/CED.2017.8308109>
- Kumala, A. R., & Wutsqa, D. U. (2016). *Perbandingan K-Means Dan Fuzzy C-Means Clustering pada Model Radial Basis Function Neural Network (Rbfn) Untuk Klasifikasi Stadium Kanker Payudara*.
- Pratama, A. Y., Prabowo, I. A., & Normassari, A. (2017). (Alpukat) Aplikasi Pengaduan Masyarakat Untuk Melaporkan Kejadian Pungutan Liar Di Kabupaten Kudus Berbasis Android. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 8(2), 399.  
<https://doi.org/10.24176/simet.v8i2.1119>
- Rante Rerung, R. (2018). Penerapan Data Mining dengan Memanfaatkan Metode Association Rule untuk Promosi Produk. *Direvisi: 27 Mei*, 3(1).  
<https://doi.org/10.31544/jtera.v3.i1.2018.89-98>
- Seberapa Pentingkah NIK, Data Kependudukan dan KTP Elektronik?* (n.d.). Retrieved February 4, 2020, from <http://dukcapil.bangka.go.id/berita/detail/seberapa-pentingkah-nik-data-kependudukan-dan-ktp-elektronik>
- Sembiring, F., & Sari, D. P. (2019). *Design Process Data Storage and Organize Data Scraping*. 1(1), 22–26.
- Windarto, A. P. (2017). Penerapan Datamining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering Method. *Techno.Com*, 16(4), 348–357.  
<https://doi.org/10.33633/tc.v16i4.1447>