



PENGLASTERAN BANK SAMPAH MENGGUNAKAN METODE K-MEANS PADA DINAS LINGKUNGAN HIDUP KABUPATEN PASURUAN

Agung Kurniawan¹⁾, Indah Dwi Mumpuni²⁾, Mohamad As'ad³⁾

Sistem Informasi, STMIK PPKIA Pradnya Paramita Malang

Email: agungkurniawan1794@gmail.com¹⁾

Email: indahstimata@yahoo.com²⁾

Email: asad.stat@yahoo.co.id³⁾

Abstrak

Sampah suatu ancaman yang serius dalam lingkungan sekitar, masalah pokoknya adalah pengolahan sampah yang kurang memadai mengakibatkan lingkungan sekitar mengalami kerusakan yang serius. Untuk mencegah kerusakan tersebut, Dinas Lingkungan Hidup memfasilitasi Bank Sampah di seluruh wilayah Kabupaten Pasuruan dalam upaya mengurangi sampah yang merusak lingkungan. Permasalahan yang sering dihadapi oleh Dinas Lingkungan Hidup adalah bagaimana metode yang efektif dan efisien dalam menentukan bank sampah yang baik dalam mengelola sampah. Tujuan penelitian ini untuk meningkatkan kinerja Dinas Lingkungan Hidup dalam menentukan anggaran dana operasional. Data yang diambil adalah 15 Bank Sampah, data tersebut dibagi menjadi 3 *dataset* yaitu jumlah nasabah, timbunan sampah dan omset. Metode yang digunakan yaitu algoritma *k-means clustering* dengan 3 kelompok. *Cluster* pertama 1 = 17; 112; 450000, *cluster* kedua 2 = 49; 275; 1018750 dan *cluster* ketiga 3 = 67; 362; 1325000. Hasilnya menunjukkan bahwa dari dua *cluster*, bernilai persis sama SSE (*Sum of Square Error*). *cluster* pertama adalah 40090915055 dan *cluster* kedua adalah 40090915055. Hasilnya dianalisis dan disimpulkan berdasarkan *cluster* yang terbentuk untuk mendukung keputusan penerimaan dana dan peralatan pengolahan sampah di dinas lingkungan kabupaten Pasuruan.

Kata kunci: algoritma K-Means, clustering, bank sampah

Abstract

Rubbish is a serious problem in the environmental, the main problem is a less of adequate rubbish processing, so the resulting in damage to the environmental seriously. To prevent damage, the department of environmental to facilitate the rubbish bank in all of part area for Pasuruan district to reduce the rubbish damaging the environmental. The problems often encountered by the Department of environmental is how effective and efficient method of determining a good rubbish bank in managing waste. The purpose of this research was to improve the performance of the Department of the environmental in determining the Fund's operational budget. The data is taken from 15 rubbish bank, the data is divided into 3 dataset namely the number of customers, pile of rubbish and turnover. The method that used namely algorithm for k-means clustering with 3 groups. The first cluster 1 = 17; 112; 450000, the second cluster 2 = 49; 275; 1018750 and the third cluster 3 = 67; 362; 1325000. The results show that from two cluster, exactly same value of the SSE (Sum of Square Error). The first cluster is 40090915055 and the second cluster is 40090915055. The results are analyzed and concluded based on the clusters formed in order to support the decision of funds acceptance and rubbish processing equipment in the department of environmental Pasuruan district.

Keywords: algorithm for K-Means, clustering, the rubbish bank



PENDAHULUAN

Sampah rumah tangga menjadi suatu ancaman yang serius untuk wilayah di Indonesia terutama di Kabupaten Pasuruan, masalah pokoknya mencakup limbah manusia dan tumpukan sampah itu sendiri, terjadinya pengolahan sampah yang kurang memadai seperti penumpukan secara tidak terkendali, pembakaran, pembuangan sampah ke sungai maupun selokan serta ke tanah kosong merupakan suatu ancaman yang paling besar di Kabupaten Pasuruan. Adapun tugas Dinas Lingkungan Hidup (DLH) Kabupaten Pasuruan sesuai peraturan Pemerintah nomor 38 tahun 2007 tentang pembagian urusan Pemerintahan diantaranya merumuskan kebijakan bidang lingkungan yang meliputi perencanaan, pengelolalaan, pemanfaatan, pengendalian, pengawasan, pemberdayaan dan penegakan hukum dalam rangka pelestarian lingkungan hidup. Pegawai Dinas Lingkungan Hidup saat ini membutuhkan penanganan data untuk mengetahui bank sampah daerah mana yang produktif maupun tidak produktif dalam mengolah sampah, serta seringkali pegawai mengalami *redundancy* data yang mengakibatkan pegawai kesulitan dalam mencari data yang di butuhkan. Dengan mengimplementasikan datamining pada Dinas Lingkungan Hidup, Maka perlu

adanya Algoritma clustering, salah satunya adalah K-Means clustering yang dapat diimplementasikan pada data bank sampah yang meliputi jumlah nasabah, jumlah timbunan sampah dan omset masing-masing bank sampah, sehingga dapat diketahui pengelompokan data sesuai bank sampah berdasarkan data tersebut. Dari masing-masing pengelompokan tersebut dapat dilihat setiap karakteristik yang diketahui kelompok mana yang dapat mengolah timbulan sampah dalam jumlah paling banyak dan sedikit.

Di sisi lain, informasi mengenai algoritma *K-Means* dapat diperoleh dari berbagai sumber. Penelitian Nasari dkk (2015) dengan judul “Penerapan *K-Means Clustering* pada Data Mahasiswa Baru (Study Kasus : Universitas Potensi Utama)” penulis dalam penelitian ini menjelaskan bahwa hasil algoritma *K-Means Clustering* yang diperoleh ada dua kelompok, pusat *cluster* dengan *cluster* 1= 1;1.75;1.5 dan *cluster* 2= 2.95;1.65;1.4 *cluster* pertama jika asal sekolah adalah SMA maka rata-rata jurusan yang diambil adalah Sistem Informasi dan kedua jika asala sekolahnya adalah SMK maka rata-rata jurusan yang diambil adalah Teknik Informatika.

Asroni dkk (2015) dengan judul “Penerapan Metode *K-Means* untuk *Clustering* Mahasiswa Berdasarkan Nilai



Akademik dengan Weka Interface Studi Kasus pada UMM Magelang” dalam penelitian ini menjelaskan bahwa algoritma *K-Means* bisa digunakan untuk mengelompokkan mahasiswa berdasarkan IPK dan beberapa atribut mata kuliah.

Metisen dkk (2015) dengan judul “Analisis *Clustering* Menggunakan Metode *K-Means* Dalam Pengelompokan Penjualan Produk pada Swalayan Fadhila.” Dalam penelitian ini menghasilkan dua jenis kelompok data yaitu data penjualan rendah dan data penjualan tinggi.

Berdasarkan permasalahan yang telah diuraikan dengan adanya metode *K-means clustering*, diharapkan membantu petugas dalam mengelompokkan bank sampah yang terbanyak menghasilkan omset pertahun di Dinas Lingkungan Hidup Kabupaten Pasuruan. Sebagai penunjang penerimaan dana maupun peralatan pengolahan sampah.

Hasil implementasi metode yang diusulkan berhasil menemukan bank sampah yang produktif dalam mengolah sampah seperti di jurnal yang terdahulu dalam mengelompokkan data *K-Means Clustering*.

KAJIAN LITERATUR

Data Mining

Data Mining adalah proses yang mempekerjakan suatu atau lebih teknik

pembelajaran komputer (*machine learning*) untuk menganalisa dan mengekstraksi pengetahuan (*knowledge*) secara otomatis. Definisi lain diantaranya adalah pembelajaran berbasis induksi (*induction-based learning*) adalah proses pembentukan definisi-definisi konsep umum yang dilakukan dengan cara mengobservasi contoh-contoh spesifik dari konsep-konsep yang akan dipelajari. *Knowledge Discovery in Database* (KDD) adalah penerapan metode saintifik pada data mining. Dalam konteks ini data mining merupakan satu langkah dari proses KDD.

Clustering

Analisis *Cluster* merupakan salah satu teknik data mining yang bertujuan untuk mengidentifikasi sekelompok obyek yang mempunyai kemiripan karakteristik tertentu yang dapat dipisahkan dengan kelompok obyek lain, sehingga obyek yang berada pada kelompok yang berbeda. Jumlah kelompok yang dapat diidentifikasi tergantung pada banyak dan variasi data obyek. Tujuan dari pengelompokan sekumpulan data data obyek kedalam beberapa kelompok yang mempunyai karakteristik tertentu dan dapat dibedakan satu sama lainnya adalah untuk analisis dan interpretasi lebih lanjut sesuai dengan tujuan penelitian yang dilakukan.



K-Means

K-Means dapat diterapkan pada data yang dipresentasikan dalam r-dimensi ruang tempat. K-Means mengelompokkan set data r-dimensi, $X = \{X_i | i=1 \dots, N\}$, di mana $X_i \in \mathbb{R}^d$ yang menandakan data ke-i sebagai “titik data”. Seperti yang dijelaskan sebelumnya bahwa K-Means partisipasi X dalam K cluster, Algoritma K-Means mengelompokkan semua titik data dalam X sehingga setiap titik X_i hanya jatuh dalam satu dari K partisi. Yang perlu diperhatikan adalah titik berada dalam cluster yang mana, dilakukan dengan cara dengan memberikan setiap titik sebuah ID cluster. Titik dengan ID cluster yang sama berarti berada dalam satu cluster yang sama, sedangkan titik dengan ID cluster yang berbeda berada dalam cluster yang berbeda untuk menyatakan hal ini, biasanya dilakukan dengan vektor keanggotaan cluster M dengan panjang N, dimana M_i bernilai ID cluster titik X_i .

Pada saat data sudah dihitung ketidakmiripan terhadap setiap centroid, maka selanjutnya dipilih ketidakmiripan yang paling kecil sebagai cluster yang akan diikuti sebagai relokasi data pada cluster disebuah iterasi. Relokasi sebuah data dalam cluster yang akan diikuti dapat dinyatakan dengan nilai keanggotaan a yang bernilai 0

atau 1. Nilai 0 juga tidak menjadi anggota cluster dan 1 jika menjadi anggota sebuah cluster. Karena K-Means mengelompokkan secara tegas data hanya pada satu cluster, maka dari nilai a sebuah data pada semua cluster, hanya satu yang bernilai 1, sedangkan lainnya 0 seperti yang dinyatakan oleh persamaan berikut:

$$a_{ij} = \begin{cases} 1 & \text{arg min}_j \{d(x_i, C_j)\} \\ 0 & \text{lainnya} \end{cases}$$

$d(x_i, C_j)$ menyatakan tidak kemiripan (jarak) dari data ke-i ke cluster C_j .

Sementara relokasi centroid untuk mendapatkan centroid C didapatkan dengan menghitung rata-rata setiap fitur dari semua data yang tergabung dalam setiap cluster. Rata-rata sebuah fitur dari semua data dalam sebuah cluster dinyatakan oleh persamaan berikut:

$$c_j = \frac{1}{N_k} \sum_{l=1}^{N_k} X_{jl}$$

N_k adalah jumlah data yang tergabung dalam sebuah cluster.

Jika diperhatikan dari langkahnya yang selalu memilih cluster terdekat, maka sebenarnya K-Means berusaha untuk meminimalkan fungsi biaya non-negatif, seperti dinyatakan oleh persamaan berikut:

$$J = \sum_{i=1}^N \sum_{l=1}^k a_{il} d(x_i, c_l)^2$$

Dengan kata lain, *K-Means* berusaha untuk meminimalkan total jarak kuadrat (*squared distance*) diantara setiap titik x_i dan representasi *cluster* c_j terdekat.

METODE PENELITIAN

Konsep sistem yang akan dibuat dalam penelitian ini adalah untuk menemukan pemecahan masalah yang sesuai dalam pengelompokan data bank sampah. Dapat dilihat pada gambar 1.1 *flowchart* Algoritma *K-Means Clustering*, alur dari algoritma dari metode *K-Means* yang digunakan dalam pengelompokkan data bank di Dinas Lingkungan Hidup Kabupaten Pasuruan di proses secara bertahap adalah sebagai berikut:

1. Menentukan titik K sebagai *centroid* awal
2. Proses menghitung *ecludian distance*
3. Pengelompokan berdasarkan *ecludian distance* terkecil
4. Memindahkan letak *centroid* ke tengah area
5. Ada pergerakan
6. Bentuk K cluster dengan menempatkan semua titik terdekat.
7. Ulangi perhitungan *centroid* dari setiap *cluster*
8. Sampai *centroid* tidak berubah
9. Hasil *K-Means*



Gambar 1.1 Flowchart Algoritma *K-Means Clustering*

Rancangan Penelitian

Data yang digunakan dalam penelitian adalah data Bank Sampah dengan periode 2 tahun terakhir yaitu tahun 2015-2016 dengan menggunakan algoritma *K-Means clustering*. Untuk mencapai tujuan tersebut, penulis melakukan pengujian dengan menggunakan data tersebut. Adapun pengelompokan data dilakukan sebagai berikut :

Sumber data dalam penelitian ini diambil dari Dinas Lingkungan Hidup Kabupaten Pasuruan dimana data kuantitatif yang terdiri atas data bank sampah 2 tahun



yaitu tahun 2015-2016 dengan banyak data 16 bank sampah.

Pada tahap selanjutnya dilakukan proses perubahan data, dengan tujuan agar data dapat diolah menggunakan algoritma *K-Means Clustering*. Data variabel yang dipilih pada data bank sampah yang diperoleh adalah jumlah nasabah, jumlah timbunan sampah dan omset

Setelah data variabel sudah ditentukan maka langkah selanjutnya adalah proses perhitungan data menggunakan algoritma *K-Means Clustering*.

HASIL PENELITIAN DAN PEMBAHASAN

Metode *K-Means* memiliki beberapa tahapan sebelum memperoleh hasil yang diinginkan sebagai berikut :

1. Data assignment

Penentuan kembali ID *cluster* dari semua titik data berdasarkan data bank sampah diantaranya jumlah nasabah, jumlah timbunan sampah dan omset. Setiap data sudah di tetapkan ke *centroid* terdekat dengan pemecahan hubungan apa adanya. Hasilnya berupa data yang terpartisi sebagai berikut:

Tabel 1.1 Data bank sampah 2 tahun terakhir mulai tahun (2015-2016)

No	Nama Bank Sampah	jn	Jts	omset
1	Langgeng Abadi	62	310	1085000
2	Berkah Sejahtera	46	257	957500
3	Bank Samsul	50	277	1027500
4	Karya Asri	51	282	1045000
5	BISS	47	265	983750
6	Panggung Jaya	51	282	1045000
7	Dahlia	52	287	1062500
8	Berkah Sejahtera II	49	275	1018750
9	Barokah	44	250	931250
10	Tejo Wangi (BST)	45	255	948750
11	Tunas Ardoma	67	362	1325000
12	Opek	57	312	1250000
13	Glamors	17	112	450000
14	Pasuruan	30	150	600000
15	Dadi mulyo	60	300	1050000

Langkah-langkah selanjutnya yang dilakukan sebagai berikut:

Inisialisasi dilakukan pemilihan K data sebagai *centroid* awal, Misalnya dipilih data ke -8, 11 dan 13.

Tabel 1.2 Data Centroid Awal

Centroid	Jn	jts	omset
1	49	275	1018750
2	67	362	1325000
3	17	112	450000

2. Relocation of "means"

Setiap representasi cluster direlokasi ke pusat (*center*) dengan rata-rata aritmatika dari semua data yang ditetapkan masuk ke dalamnya. Rasionalnya langkah ini



didasarkan pada observasi bahwa dalam memberikan set titik, representasi tunggal yang terbaik untuk set tersebut (dalam hal meminimalkan jumlah kuadrat jarak *Euclidean* di antara setiap titik data *representatif*) adalah dari rata-rata dari titik data. Algoritma *K-Means* mencapai kondisi *konvergen* ketika pengalokasian kembali titik data (dan juga lokasi *centroid*) tidak lagi berubah. Proses iterasi ke iterasi hingga dicapai kondisi *konvergen* juga dapat diamati dari nilai fungsi *objektif* yang didapatkan.

Pemilihan *K* titik data sebagai *centroid* awal juga memengaruhi hasil *clustering*. Sifat ini menjadi karakteristik alami *K-Means* yang dapat mengakibatkan hasil *cluster* yang didapat pada percobaan berbeda mendapatkan hasil yang berbeda juga. Kondisi ini dikenal dengan solusi *local optima*, yang artinya algoritma *K-Means* sangat *sensitive* terhadap alokasi awal *centroid*. Penyelesaian masalah *local optima* dapat diselesaikan dengan menjalankan algoritma beberapa kali dengan inisial *centroid* yang berbeda kemudian memilih hasil yang terbaik. Berikut data bank sampah yang lakukan dengan metode *K-Means* :

- Step pertama iterasi 1

Menghitung jarak setiap data ke *centroid* terdekat. *Centroid* terdekat akan menjadi *cluster* yang diikuti oleh data tersebut.

Berikut contoh perhitungan jarak ke setiap *centroid* pada data ke -1:

$$d(x_1, c_1) = \sum_{i=1}^r (x_{1i} - c_{1i})^2 = (17-62)^2 + (112-310)^2 + (450000-1085000)^2 = 635,000$$

$$d(x_1, c_2) = \sum_{i=1}^r (x_{1i} - c_{2i})^2 = (49-62)^2 + (275-310)^2 + (1018750-1085000)^2 = 66,250$$

$$d(x_1, c_3) = \sum_{i=1}^r (x_{1i} - c_{3i})^2 = (67-62)^2 + (362-310)^2 + (1325000-1085000)^2 = 240,000$$

Tabel 1.3 Proses Iterasi ke-1

No	CENTROID 1			CENTROID 2			CENTROID 3		
	17	112	450000	49	275	1018750	67	362	1325000
1	635,000.03			66,250.01			240,000.01		
2	507,500.02			61,250.00			367,500.02		
3	577,500.02			8,750.00			297,500.01		
4	595,000.03			26,250.00			280,000.01		
5	533,750.02			35,000.00			341,250.01		
6	595,000.03			26,250.00			280,000.01		
7	612,500.03			43,750.00			262,500.01		
8	568,750.02			-			306,250.01		
9	481,250.02			87,500.00			393,750.02		
10	498,750.02			70,000.00			376,250.02		
11	875,000.04			306,250.01			-		
12	800,000.03			231,250.00			75,000.02		
13	-			568,750.02			875,000.04		
14	150,000.01			418,750.02			725,000.03		
15	600,000.03			31,250.01			275,000.01		

Jarak terdekat berwarna merah dengan *cluster* yang diikuti

Tabel 1.4 Hasil Iterasi ke-1

No	CLUSTER 1			CLUSTER 2			CLUSTER 3		
1				62	310	1085000			
2				46	257	957500			
3				50	277	1027500			
4				51	282	1045000			
5				47	265	983750			
6				51	282	1045000			
7				52	287	1062500			
8				49	275	1018750			
9				44	250	931250			
10				45	255	948750			



11							67	36 2	13 25 00 0
12							57	31 2	12 50 00 0
13	17	11 2	45 00 00						
14	30	15 0	60 00 00						
15				60	300	105000 0			

Menghasilkan *centroid* baru diantaranya sebagai berikut:

Tabel 1.5 Data Centroid Baru

Centroid	jn	jts	Omset
1	23.5	131	525000
2	50.63636	276.3636	1014090.9
3	62	337	1287500

Hasil *centroid* baru di hitung dengan data *cluster* awal berikut proses perhitungannya :

$$(23.5 - 17)^2 + (131 - 112)^2 + (52500 - 45000)^2 = 5.625.000.403$$

$$(50.636 - 62)^2 + (276.36 - 310)^2 + (1014091 - 1085000)^2 = 5.028.100.434$$

$$(62 - 67)^2 + (337 - 362)^2 + (1287500 - 1325000)^2 = 1.406.250.650$$

Didapatkan nilai (*Sum of Square Error*) SSE = 40090915055

Karena perubahan fungsi SSE masih diatas ambang batas yang ditetapkan, maka proses dilanjutkan ke iterasi selanjutnya.

- Step kedua iterasi 2

Menghitung jarak setiap data ke *centroid* terdekat. *Centroid* terdekat akan menjadi *cluster* yang diikuti oleh data tersebut.

Tabel 1.6 Proses Iterasi ke-2

No	CENTROID 1			CENTROID 2			CENTROID 3		
	24	131	525000	51	276	1014091	62	337	1287500
1	560,000.03			70,909.10			202,500.00		
2	432,500.02			56,590.91			330,000.01		
3	502,500.02			13,409.09			260,000.01		
4	520,000.02			30,909.09			242,500.01		
5	458,750.02			30,340.91			303,750.01		
6	520,000.02			30,909.09			242,500.01		
7	537,500.02			48,409.09			225,000.01		
8	493,750.02			4,659.09			268,750.01		
9	406,250.02			82,840.91			356,250.01		
10	423,750.02			65,340.91			338,750.01		
11	800,000.03			310,909.10			37,500.01		
12	725,000.02			235,909.09			37,500.01		
13	75,000.00			564,090.93			837,500.03		
14	75,000.00			414,090.93			687,500.03		
15	525,000.03			35,909.10			237,500.00		

Jarak terdekat berwarna merah dengan diikuti *cluster* sebagai berikut :

Tabel 1.7 Hasil Iterasi ke-2

n o	CLUSTER 1			CLUSTER 2			CLUSTER 3		
	1				62	310	1085 000		
2				46	257	9575 00			
3				50	277	1027 500			
4				51	282	1045 000			
5				47	265	9837 50			
6				51	282	1045 000			
7				52	287	1062 500			
8				49	275	1018 750			
9				44	250	9312 50			
10				45	255	9487 50			
11							67	362	132 500 0
12							57	312	125 000 0
13	17	11 2	45 00 00						
14	30	15 0	60 00						



			00					
1			60	300	1050			
5					000			

Menghasilkan *centroid* baru diantaranya sebagai berikut :

Tabel 1.8 Data Centroid Akhir

Centroid	jn	Jts	Omset
1	23.5	131	525000
2	50.63636	276.3636	1014090.9
3	62	337	1287500

Hasil *centroid* baru di hitung dengan data *cluster* awal berikut proses perhitungannya :

$$(23.5 - 17)^2 + (131 - 112)^2 + (52500 - 450000)^2 = 5.625.000.403$$

$$(50.636 - 62)^2 + (276.36 - 310)^2 + (1014091 - 1085000)^2 = 5.028.100.434$$

$$(62 - 67)^2 + (337 - 362)^2 + (1287500 - 1325000)^2 = 1.406.250.650$$

Didapatkan nilai (*Sum of Square Error*) **SSE = 40090915055**

Karena perubahan nilai SSE sama dengan nilai SSE iterasi yang pertama sama, berarti kondisi *cluster* sudah mencapai *konvergen* dan proses iterasi pun berhenti.

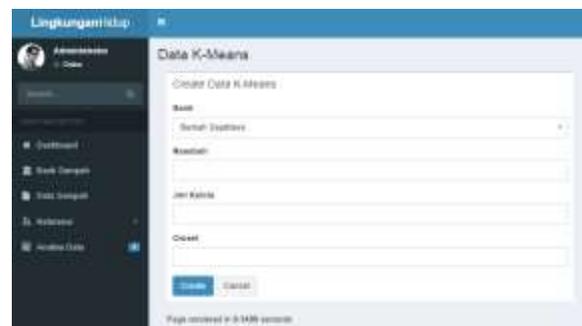
IMPLEMENTASI SISTEM

Pada proses implementasi sistem datamining ini, dibuat dengan menggunakan bahasa pemrograman berbasis *web*. Pada sistem ini dirancang menggunakan algoritma *K-Means Clustering*. Proses yang dilakukan pertama-tama pilih menu analisa data, pada

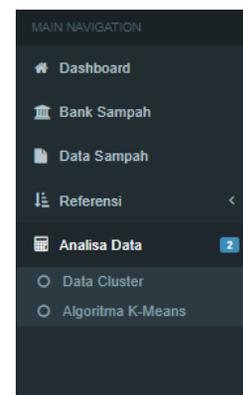
menu analisa data terbagi menjadi 2 sub menu yaitu menu data cluster dan Algoritma K-Means lihat **gambar 1.4**. pada menu data cluster akan menampilkan data pada **gambar 1.2**. Untuk proses *input* data pilih menu Data cluster maka akan tampil pada **gambar 1.3**, data yang diinputkan merupakan data yang akan diproses melalui algoritma *K-Means Clustering*.



Gambar 1.2 Halaman Data Cluster



Gambar 1.3 Halaman Form Input Data



Gambar 1.4 Menu Analisa Data



Langkah selanjutnya setelah data sudah diinputkan. Pilih menu Algoritma K-Means maka akan tampil data secara keseluruhan dapat dilihat pada **gambar 1.5**.

No	Nama Bank	Jumlah Sampah	Jumlah Tersebut Bank	Jumlah
1	Lampung Barat	52	178	100000
2	Bank Lampung	50	207	107500
3	Bank Sumsel	50	277	117500
4	Maya Bank	50	300	120000
5	BBK	50	300	120000
6	Pengadilan Jari	50	300	120000
7	Bank	50	307	122500
8	Bank Lampung 2	50	275	110000
9	Bank	50	200	80000
10	Pop Wang (BKI)	50	50	10000
11	Bank-Bank	50	100	100000
12	Bank	50	177	70000
13	Bank	50	100	10000
14	Bank	50	100	10000
15	Bank	50	100	10000

Gambar 1.5 Data Cluster

Setelah tampil data selanjutnya pilih menu proses data maka akan tampil *Form* pada **gambar 1.6**. pada *form* tersebut pilih *centroid* di setiap *cluster* atau bisa dilakukan secara random maupun input data secara manual.

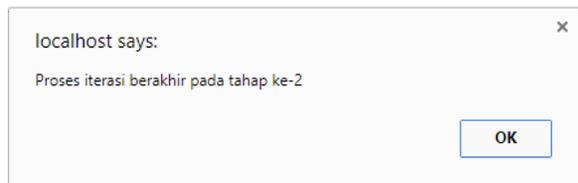
Gambar 1.6 Input Data Centroid

Setelah *form* terisi semua langkah selanjutnya klik tombol proses maka akan tampil **gambar 1.7**. Pada gambar tersebut hasil proses perhitungan algoritma K-Means *clustering* iterasi yang ke-1.

Gambar 1.7 Halaman Iterasi ke-1

Proses iterasi ke-1 selesai klik tombol proses iterasi selanjutnya hingga proses iterasi selesai dapat dilihat pada **gambar 1.8** dan **gambar 1.9**

Gambar 1.8 Halaman Iterasi ke-2



Gambar 1.9 Halaman Proses Iterasi Selesai

Gambar 1.10 Hasil Iterasi Algoritma K-Means



Proses perhitungan selesai dan akan tampil gambar 1.10 hasil iterasi algoritma *K-Means* maka dapat disimpulkan hasil iterasi yang terbentuk terbagi menjadi 3 kelompok dengan *cluster* pertama ada 2 bank sampah, *cluster* kedua 11 bank sampah dan *cluster* ketiga 2 bank sampah.

KESIMPULAN DAN SARAN

Berdasarkan penelitian yang telah dilakukan dapat disimpulkan bahwa pengklasteran bank sampah menggunakan metode *K-Means* ini dapat disimpulkan sebagai berikut:

1. Memudahkan admin untuk melihat data sesuai kelompok yang ditentukan dalam pemberian dana operasional yang layak ke setiap bank sampah.
2. Pengelompokan data ditentukan pusat *cluster* 1 = 17;112;450000, *cluster* 2 = 49;275;1018750 dan *cluster* 3 = 67;362;1325000, kemudian dilakukan perhitungan algoritma *K-Means* dengan data yang ditentukan. Hasilnya menunjukkan bahwa dari dua cluster bernilai SSE (*Sum of Square Error*) sama yaitu *cluster* pertama adalah 40090915055 dan *cluster* kedua adalah 40090915055
3. Pengambilan keputusan dana operasional dan peralatan pengolahan

sampah diambil dari pengelompokan data *cluster* 3 dari iterasi yang ke 2 yaitu bank Sampah Tunas Ardoma dengan nilai *cluster* 67; 362; 1325000 dan Bank Sampah Opek dengan nilai *cluster* 57; 312; 1250000.

Adapun saran yang dapat disampaikan untuk peningkatan dan pengembangan sistem pengklasteran bank sampah menggunakan metode *K-Means* di Dinas Lingkungan Hidup Kabupaten Pasuruan adalah adanya algoritma lain untuk bahan perbandingan dari penelitian yang sudah dilakukan.

REFERENSI

- Asroni, Adrian R. 2015. *Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang*. Vol.18, No. 1, hal 76-82
- Hermawati. 2013. *Data Mining*. Yogyakarta: CV ANDI OFFSET
- Metisen B M, Sari H L. 2015. *Analisis Clustering Menggunakan Metode K-Means Dalam Pengelompokan Penjualan Produk Pada Swalayan Fadhila*. Vol. 11 No. 2. ISSN ; 1858-2680
- Nasari F, Darma S.2015. *Penerapan K-Means Clustering Pada Data Penerimaan Mahasiswa Baru Studi Kasus Universitas Potensi Utama*. ISSN : 2302-3805
- Nofriansyah, Nurcahyo. 2015. *Algoritma Data Mining dan Pengujian*. Yogyakarta: Deepublish



Prasetyo. 2014. *DATA MINING – Mengolah
Data menjadi Informasi
Menggunakan Matlab*. Yogyakarta:
Penerbit ANDI

Ratner Bruce. 2011. *Statistical and Machine-
Learning Data Mining*. Boca Raton:
CRC Press