

CLUSTERING OBJECT RETRIBUSI BERBASIS K-MEANS

CLUSTERING OBJECT RETRIBUTION BASE K-MEANS

Ertina Sabarita Barus¹, Bersama Sinuraya¹, dan Jenni Veronika Ginting¹

¹STMIK Kristen Neumann Indonesia
Jl. Letjen Jamin Ginting KM. 10,5 Medan
baruschild@gmail.com

Program Studi Teknik Informatika

ABSTRAK

. Pendapatan daerah merupakan penerimaan dana bagi pemerintahan daerah yang digunakan sebagai penunjang pembangunan daerah. Pendapatan daerah digunakan untuk membiayai proyek-proyek, program-program pemerintah dan kegiatan-kegiatan daerah, salah satu pendapatan daerah yaitu retribusi daerah, ada beberapa jenis retribusi daerah dan masing-masing jenis retribusi tersebut terdiri dari titik-titik objek retribusi. Daerah-daerah yang menjadi objek retribusi sudah ditentukan oleh pemerintahan daerah. Kegiatan transaksi pembayaran retribusi dilakukan secara langsung oleh petugas pemungut retribusi dan wajib retribusi. Hal ini rentan menyebabkan terjadi pemungutan liar yang dilakukan oleh oknum pemungut retribusi. Karena tidak ada pengawasan dan monitoring yang dilakukan oleh atasan. Hal tersebut sulit dilakukan karena titik-titik objek retribusi yang jumlahnya sangat banyak di seluruh wilayah pemerintahan daerah karo. Pada penelitian ini akan dilakukan riset untuk mendefinisikan seluruh titik-titik retribusi kemudian dilakukan clustering. Clustering berbasis K-Means

Kata Kunci: Clustering, K-Means, object retribusi.

PENDAHULUAN

Objek Retribusi di daerah kabupaten karo mejadi salah satu retribusi yang memberikan pendapatan yang besar disebabkan kabupaten karo memiliki titik –titik objek retribusi yang tersebar luas didearah kabupaten karo yang luasnya 2.216 Km dengan jumlah penduduk sebanyak 358.823 sumber data Pariwisatasumut.net. kabupaten karo memiliki 10 kota yang tersebar di 17 kecamatan diantaranya adalah kabanjahe, berastagi, tiga panah, barus jahe, tiga binanga, munthe, simpang empat, merek, lau baleng dan mardinding.

Di beberapa titik pemungutan wajib parkir masih sering terjadi juru parkir yang meminta imbbalana secara tidak langsung pada lokasi parkir. Dan masih kurangnya pengawasan yang dilakukan oleh dinas terkait masih kurang karena masih ada para pengguna lahan parkir yang tidak tahu harus kemana jika ingin menyampaikan kritikan (Reski, Puspasari, 2015). Kondisi pengawasan yang sangat minim memungkinkan untuk oknum petugas pemungut retribusi melalukan pemungutan liar yang tentunya hal ini merugikan pemerintah daerah. Kegiatan tersebut juga berdampak bagi wajib retribusi merasa tidak nyaman karena

merasa dana retribusi yang diberikan digunakan tidak sesuai dengan semestinya. Selain itu oknum pungli akan mendapatkan pidana jika hal ini sudah terbukti dilakukan penyelewengan dana retribusi.

Analisis cluster adalah sebuah metodologi untuk mengklasifikasi sampel-sampel secara otomatis kedalam sejumlah kelompok-kelompok dengan nilai tertentu, sehingga beberapa sampel yang sama berkumpul didalam sebuah kelompok dan terpisah dengan kelompok lain dengan nilai sampel yang berbeda. (kantardzic, 2003). Yang menjadi input dari sebuah sistem analisis cluster adalah sebuah set sampel dan sebuah ukuran kesamaan ataupun ketidaksamaan diantara sampel. Sedangkan output dari sebuah sistem cluster analisis yaitu sejumlah kelompok (cluster) yang sudah terbagi didalam sebuah partisi-partisi ataupun struktur sebuah partisi. Sampel-sampel didalam sebuah clustering merupakan perwakilan dari pengukuran sebuah vector atau titik-titik ruang multidimensional. Clustering merupakan sebuah masalah yang sulit karena data dapat membangkitkan sebuah clustering dengan perbedaan bentuk dan ukuran pada sebuah ruang n-dimensional.

Steinbach et al metode clustering melakukan penelitian dengan mengelompokkan dokumen-dokumen dengan membandingkan metode K-Means yang standart dengan bisecting K-Means algoritma. Untuk pengelompokan dokumen pendekatan metode bisecting K-Means baik diterapkan. (steinbach,) Morissette dalam penelitiannya melakukan clustering data pada database yang besar dengan multidimensidata dengan melakukan pendekatan 3 algoritma K_menas yaitu Forgy Lloyd algoritma, MacQueen Algoritma dan Hartigan dan

Wong algoritma menghasilkan beberapa model matematika dan ilustrasi teknik aplikasi (Morissette, 2013), (Wahyuni, 2007) Dalam penelitiannya melakukan cluster data dengan jumlah yang sangat besar dengan menerapkan aturan similarity dan dissimilarity antara sampel set yang tersedia. Penelitian Hamerly membandingkan algoritma K-Means Standar dengan beberapa alternatif algoritma.

METODE PENELITIAN

Setelah objek-objek retribusi di petakan kedalam peta digital kemudian di ekspor ke dalam aplikasi selanjutnya dilakukan proses pengelompokkan berdasarkan nilai kesamaan dan ketidaksamaan antara sampel. Berikut terdapat N sampel dengan n dimensional dan akan dibagi kedalam beberapa partisi yaitu K cluster (C1, C2,Ck) masing-masing Ck dan nk tepat pada sebuah cluster sehingga $\sum nk = N$, dimana k =1,, K. dimana nilai rata-rata Mk dalam cluster Ck didefinisikan sebagai sebuah centroid dalam sebuah cluster dan dihitung dengan rumus :

$$M_k = \bar{X}_k = \left(\frac{1}{n_k}\right) \sum_{i=1}^{n_k} X_{ik}$$

Pada persamaan diatas dimana X_{ik} adalah nilai I yang dimiliki oleh sebuah cluster Ck. Dan akar error akar pangkat cluster Ck adalah jumlah kuadrat dari Euclidean antara Ck dan centroid. Nilai error ini disebut dengan variasi cluster

$$e_k^2 = \sum_{i=1}^{n_k} (x_{ik} - M_k)^2$$

Akar pangkat error untuk seluruh cluster K adalah jumlah sigma dari seluruh variasi cluster dan dihitung dengan menggunakan rumus berikut ini.'

$$E_k^2 = \sum_{i=1}^K e_k^2$$

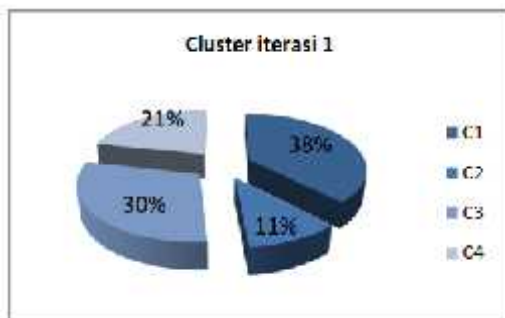
Nilai dari akar pangkat error merupakan metode pengelompokan untuk menemukan isi dari sebuah partisi cluster. Cluster K merupakan nilai minimal E_k^2 untuk sebuah nilai K. Menggunakan metode K-Means maka dilakukan pengelompokan (cluster) terhadap seluruh objek-objek wisata. Langkah-langkah pengelompokan dilakukan secara terus menerus sampai semua titik titik sampel objek retribusi masuk ke dalam sebuah cluster.

HASIL DAN PEMBAHASAN

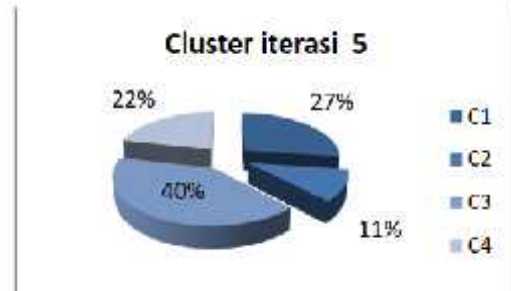
Proses K-Means clustering akan terus melakukan iterasi hingga data hasil clustering akan sama dengan hasil iterasi sebelumnya. Proses ini dilakukan berulang ulang dengan menggunakan metode penentuan centroid hingga hasil clustering pada iterasi tersebut akan menghasilkan clustering yang persis sama dengan hasil clustering pada iterasi sebelumnya.

Proses selanjutnya adalah clustering pada iterasi ke 2, yang hasil cluster belum sama dengan iterasi 1. Selanjutnya dilakukan iterasi ke 3 yang hasilnya masih ada perbedaan pada iterasi ke 2 dan begitu seterusnya hingga didapatkan hasil yang sesuai.

Pada iterasi ke 5 akhirnya didapatkan clustering yang sama dan sesuai dengan hasil pada iterasi ke 4. Berikut ditampilkan data hasil clustering dan grafik hasil clustering pada iterasi ke 5



Proses K-Means Clustering pada iterasi 1 sudah didapatkan sesuai dengan tabel 3. Data yang sudah dikelompokkan pada iterasi 1 adalah hasil dari proses clustering dimana hasil iterasi 1 dapat juga dilihat pada gambar di bawah ini:



KESIMPULAN

Untuk melakukan pengelompokan atas transaksi yang terjadi pada perusahaan dapat dilakukan dengan menerapkan K-Means Clustering. Data data yang diperoleh, diproses dengan menggunakan software Microsoft Excel ataupun dengan menggunakan software lainnya seperti SPSS ataupun Rapidminer. Nilai centroid ditentukan dengan membagi menjadi 4 kelompok

Clustering berdasarkan letak geografis terbagi 4 yaitu cluster 1 dan cluster 2. dan pajak berupa hotel yaitu cluster 3 dan penginapan cluster 4. Dari proses interasi ke 5 terdapat cilia c3 sebesar 40% dan cluster C4 yang paling kecil yaitu 11%

DAFTAR PUSTAKA

- [1] E. Kriswati, Deformasi Gunung Api Bromo pada Peningkatan Aktivis Vulkanik 2006-2007, *Widyariset*, 11(1), 2008, pp. 27-36.
- [2] R. Medaglia, eParticipation research: Moving characterization forward (2006–2011), *Government Information Quarterly*, vol. 29, no. 3, 2012, pp. 346–360.

- [3] Michael Steinbach, George Karypis dan Vipin Kumar, tahun, “ A comparison of Document Clustering Techniques”, Minneapolis,
- [4] Laurence Morissette dan Sylvain Chartier, 2013, “ The K-Means Clustering Technique: General considerations and implementation in mathematica”, Tutorial in Quantitative methods or Psychology 2013, vol. 9(1), P.15-24
- [5] Sri Andayani, 2007, “ Pembentukan cluster dalam Knowledge Discovery in Database dengan Algoritma K-Means” SEMNA MAtematika dan Pendidikan matematika 2007, Yogyakarta
- [6] Greg Hamerly dan Charles Elkan, 2002, “ alternative to K-Means Algorithm that find better clustering”jurnal apa
- [7] Bruce Eckel, 2006, Thinking in Java, Massachusetts, Prentice Hall
- [8] Ardeana Reski dan Devi Pusposari , 2016, Analisis Optimalisasi Retribusi di Tepi Jalan Umum di Kabupaten Tulungagung, Jurnal Ilmiah Mahasiswa FEB vol 4 No 2, Universitas Brawijaya.
- [9] Eko Budiyanto, 2010, Sistem Informasi Geografis dengan ArcView GIS, Andi, Yogyakarta.