

PENERAPAN DATAMINING DALAM MENGELOMPOKKAN KUNJUNGAN WISATAWAN KE OBJEK WISATA UNGGULAN DI PROV. DKI JAKARTA DENGAN K-MEANS

Linda Maulida

Program Studi Manajemen Informatika, AMIK BSI Tangerang
Komplek BSD Sektor XIV-C11 Jl Letnan Sutopo Banten, Indonesia e-mail : linda.lma@bsi.ac.id

Abstract

Visits of foreign tourists to Indonesia can increase the country's foreign exchange and improve the economy of people in tourist areas. Jakarta is the capital of Indonesia which became one tourist destination for tourists. There are 8 leading tourist attraction in DKI Jakarta province according to BPS data of Prov. DKI Jakarta 1) Taman Impian Jaya Ancol, 2) Taman Mini Indonesia Indah, 3) Kebon Binatang Ragunan, 4) National Monument, 5) National Museum, 6) Satria Mandala Museum, 7) Jakarta History Museum And 8) Sunda Kelapa Harbor . The purpose of this study is to analyze the application of datamining in classifying the number of foreign tourists visiting the Prov. DKI Jakarta using k-means. The source of research data is from BPS Prov. DKI Jakarta. The research data used is the number of tourist visitors in 2007-2013 in accordance with BPS Prov. DKI Jakarta. The data are grouped into 3 clusters namely C1 = the number of high tourist visits, C2 = the number of tourists visiting medium and C3 = the number of tourist visits is low. The final centroid value used at C1 = 15.438.488, C2 = 4.464.577 and C3 = 342.332. So that the result of grouping C1 = Taman Impian Jaya Ancol, C2 = Taman Mini Indonesia Indah Dan Kebon Binatang Ragunan and C3 = National Monument, National Museum, Satria Mandala Museum, Jakarta History Museum and Sunda Kelapa Harbor. The result of C3 grouping becomes a record for the government of Prov. DKI. Jakarta.

Keywords : Datamining, K-Means, Tourist Attraction, Jakarta

Abstrak

Kunjungan wisatawan asing ke indonesia dapat meningkatkan devisa negara dan meningkatkan perekonomian masyarakat di daerah wisata. Jakarta adalah ibukota indonesia yang menjadi salah satu tujuan wisata bagi turis. Ada 8 objek wisata unggulan di provinsi DKI jakarta menurut data BPS Prov. DKI Jakarta yakni 1) Taman Impian Jaya Ancol, 2) Taman Mini Indonesia Indah, 3) Kebon Binatang Ragunan, 4) Monumen Nasional, 5) Museum Nasional, 6) Museum Satria Mandala, 7) Museum Sejarah Jakarta Dan 8) Pelabuhan Sunda Kelapa. Tujuan penelitian ini adalah melakukan analisis penerapan datamining dalam mengelompokkan jumlah kunjungan wisatawan asing ke Prov. DKI Jakarta menggunakan k-means. Sumber data penelitian berasal dari BPS Prov. DKI Jakarta. Data penelitian yang digunakan adalah jumlah pengunjung wisatawan tahun 2007-2013 sesuai dengan BPS Prov. DKI Jakarta. Data dikelompokkan menjadi 3 cluster yaitu C1= jumlah kunjungan wisatawan tinggi, C2=jumlah kunjungan wisatawan sedang dan C3= jumlah kunjungan wisatawan rendah. Nilai centroid akhir yang digunakan pada C1= 15.438.488, C2= 4.464.577 dan C3= 342.332. Sehingga diperoleh hasil pengelompokan C1= Taman Impian Jaya Ancol, C2= Taman Mini Indonesia Indah Dan Kebon Binatang Ragunan dan C3= Monumen Nasional, Museum Nasional, Museum Satria Mandala, Museum Sejarah Jakarta Dan Pelabuhan Sunda Kelapa. Hasil pengelompokan C3 menjadi catatan bagi pemerintah Prov. DKI. Jakarta.

Kata Kunci : Datamining, K-Means, Objek Wisata, Jakarta

1. PENDAHULUAN

Kunjungan wisatawan asing ke Indonesia akan meningkatkan devisa negara serta dapat meningkatkan perekonomian masyarakat di daerah wisata. Jakarta adalah ibukota Indonesia dimana menjadi salah satu objek tujuan wisata bagi turis asing untuk datang ke Indonesia. Sejak pertengahan September 2017 dan masih berlangsung sampai hari ini. Badan Pusat Statistik (BPS) pusat mencatat kedatangan wisatawan mancanegara (wisman) sepanjang waktu tersebut trennya terus menurun. Jika dilihat dari rata-rata kunjungan wisatawan tahun 2017-2013. Khususnya kunjungan wisatawan ke objek wisata unggulan di Prov. DKI Jakarta. Tercatat ada 8 (delapan) objek wisata unggulan yang selalu dikunjungi wisatawan di Prov. DKI Jakarta (sumber BPS). Objek wisata unggulan tersebut adalah Taman Impian Jaya Ancol, Taman Mini Indonesia Indah, Kebon Binatang Ragunan, Monumen Nasional, Museum Nasional, Museum Satria Mandala, Museum Sejarah Jakarta Dan Pelabuhan Sunda Kelapa. Keberhasilan pembangunan Pariwisata ditandai dengan meningkatnya arus kunjungan Wisatawan Nusantara (Wisnus) dan wisatawan Mancanegara (Wisman) akan memberikan manfaat kesejahteraan bagi masyarakat luas, bagi usaha Pariwisata dan usaha terkait serta Pemerintah selaku Pembina Pariwisata di daerah.

Larose dalam buku yang ditulis oleh Kusri dan Luthfi mengelompokkan Data Mining dapat dibagi menjadi 6 kelompok yaitu deskripsi, estimasi, prediksi, klasifikasi, clustering (pengelompokan), dan asosiasi Kusri [2009]. Clustering akan melakukan pengelompokan data-data ke dalam sejumlah kelompok (cluster) berdasarkan kesamaan karakteristik masing-masing data pada kelompok-kelompok yang ada [Prasetyo, 2012]. Banyak metode yang bisa digunakan untuk melakukan clustering diantaranya : metode K-Means, metode LVQ (Learning Vector Quantization), FCM (Fuzzy C-Means), dan lain sebagainya. Dalam melakukan penelitian ini, penulis tidak terlepas dari penelitian-penelitian sebelumnya yang beserta metode-metode yang digunakan. Penelitian-penelitian sebelumnya merupakan acuan bagi penulis untuk melakukan dan membuat penelitian ini. Dimana penelitian-penelitian sebelumnya merupakan bahan perbandingan untuk mengetahui kekurangan dan kelebihan serta mengetahui dan membandingkan metode-metode yang digunakan dalam penelitian sebelumnya. Penulis menggunakan 2 penelitian terdahulu yang berkaitan dengan objek wisata yakni Masruro et al. [2014] dan Wahyu and Anggriawan [2015]. Penelitian 1) [Masruro et al., 2014] tentang Sistem Penunjang Keputusan Penentuan Lokasi Wisata Menggunakan K-Means Clustering Dan Topsis. Kriteria yang digunakan dalam penelitian (tema, fasilitas dan budget). Tujuan penelitian ini adalah kolaborasi metode TOPSIS dan K-Means untuk menghasilkan sebuah informasi berupa daftar tempat wisata yang sesuai dengan tingkat kepentingan yang diharapkan oleh pengguna dari kriteria yang telah ditentukan terlebih dahulu dan 2) [Wahyu and Anggriawan, 2015] tentang Sistem Rekomendasi Paket Wisata Se-Malang Raya Menggunakan Metode Hybrid Content Based Dan Collaborative. Tujuan penelitian ini merekomendasi suatu dengan menggunakan data history transaksi yang diproses dengan metode Content-based dan metode Collaborative, selanjutnya dilakukan proses hybrid dan menghasilkan hasil rekomendasi paket wisata yang sesuai bagi wisatawan.

Berdasarkan penelitian terdahulu, penelitian yang dilakukan peneliti tentang pengelompokan objek wisata unggul berbeda. Dimana perbedaannya terletak pada objek penelitian yang diteliti. Tujuannya adalah untuk mengetahui objek wisata unggulan di Prov. DKI Jakarta yang memiliki potensi yang paling rendah dalam kunjungan wisatawan ke Indonesia khususnya Prov. DKI Jakarta. Pengelompokan tersebut dapat menggunakan metode pengelompokan dengan algoritma K-Means. Dalam hal ini, penerapan datamining mampu menjadi solusi dengan menganalisa data. Perlu diketahui bahwa datamining merupakan suatu alat yang memungkinkan para pengguna untuk mengakses secara cepat data dengan jumlah yang besar. Pengertian yang lebih khusus lagi dari data mining, yaitu suatu alat dan aplikasi menggunakan analisis statistik pada data. Data mining juga dikenal sebagai Knowledge Data Discovery di dalam basis data. Data mining adalah suatu proses ekstraksi atau penggalian data dan informasi yang besar, yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari database yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting [Soni and Ganatra, 2012]. Dengan data yang sudah dikelompokkan menggunakan algoritma K-Means diharapkan dapat mempermudah Prov. DKI Jakarta dalam mengelompokkan hasil jumlah kunjungan wisatawan ke objek wisata unggulan yang kunjungan wisatanya paling sedikit, sehingga Pemerintahan Prov. DKI Jakarta dapat melakukan perbaikan sarana dan prasarana objek wisatawan unggulan

yang dapat meningkatkan jumlah kunjungan wisatawan yang berdampak pada pengenalan objek wisata dan peningkatan devisa negara.

2. METODE PENELITIAN

Dalam penelitian ini metode yang digunakan yaitu metode data mining sebagai berikut. (a) Tahap pengumpulan data, (b) Tahap pengolahan data, (c) Tahap Clustering dan (d) Tahap Analisis. Gambar 1 berikut adalah flowchart dalam menentukan cluster dengan K-Means [Windarto, 2017].



Gambar 1 : Flowchart K-Means
Sumber : [Windarto, 2017]

2.1 Data Mining

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistic, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [Afrisawati, 2013].

Data mining juga merupakan metode yang digunakan dalam pengolahan data berskala besar oleh karena itu data mining memiliki peranan yang sangat penting dalam beberapa bidang kehidupan diantaranya yaitu bidang industri, bidang keuangan, cuaca, ilmu dan teknologi. Dalam data mining juga terdapat metode metode yang dapat digunakan seperti klasifikasi, clustering, regresi, seleksi variabel, dan market basket analisis [Atthina and Iswari, 2014].

2.2 Clustering

Analisis Pengelompokan/ Clustering merupakan proses membagi data dalam suatu himpunan ke dalam beberapa kelompok yang kesamaan datanya dalam suatu kelompok lebih besar dari-

pada kesamaan data tersebut dengan data dalam kelompok lain [Windarto, 2017]. Potensi clustering adalah dapat digunakan untuk mengetahui struktur dalam data yang dapat dipakai lebih lanjut dalam berbagai aplikasi secara luas seperti klasifikasi, pengolahan gambar, dan pengenalan pola [Ong, 2013].

Pada proses analisis cluster metode yang digunakan untuk membagi data menjadi subset data berdasarkan kesamaan atau kemiripan yang telah ditentukan sebelumnya. Jadi analisis cluster secara umum dapat dikatakan bahwa [Wardhani, 2016]:

- a. Data yang terdapat dalam satu cluster memiliki tingkat kesamaan yang tinggi, dan
- b. Dan yang terdapat dalam suatu cluster yang berbeda memiliki tingkat kesamaan yang rendah

2.3 Algoritma K-Means

K-means merupakan salah satu metode pengelompokan data nonhierarki (sekatan) yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan ke dalam kelompok yang lain. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diatur dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok [Suprihatin, 2011].

Algoritma K-means pada dasarnya melakukan 2 proses yakni proses pendeteksian lokasi pusat cluster dan proses pencarian anggota dari tiap-tiap cluster. Proses clustering dimulai dengan mengidentifikasi data yang akan dikluster, X_{ij} ($i=1, \dots, n$; $j=1, \dots, m$) dengan n adalah jumlah data yang akan dikluster dan m adalah jumlah variabel. Pada awal iterasi, pusat setiap cluster ditetapkan secara bebas (sembarang), C_{kj} ($k=1, \dots, k$; $j=1, \dots, m$). Kemudian dihitung jarak antara setiap data dengan setiap pusat cluster [Soni and Ganatra, 2012]. Proses dasar algoritma k-means dapat dilihat di bawah ini :

1. Tentukan jumlah kluster yang ingin dibentuk dan tetapkan pusat cluster k .
2. Menggunakan jarak euclidean kemudian hitung setiap data ke pusat cluster

$$d(i, k) = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2}. \quad (1)$$

3. Kelompokkan data ke dalam cluster dengan jarak yang paling pendek dengan persamaan

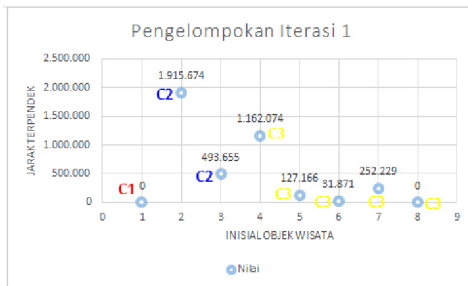
$$\min_k \sum_i^i -a_{ik} = \sqrt{\sum_j^m (C_{ij} - C_{kj})^2} \quad (2)$$

4. Hitung pusat cluster yang baru menggunakan persamaan

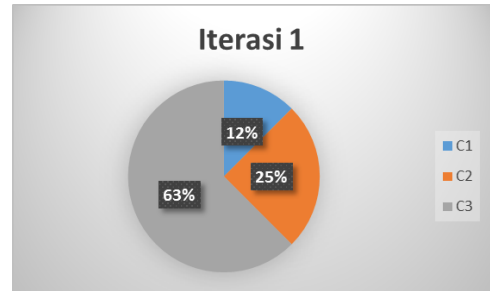
$$C_{kj} = \frac{\sum_k^i x_{ij}}{p} \quad (3)$$

Dengan : $X_{ij} \in$ Kluster ke k $p =$ banyaknya anggota kluster ke - k

5. Ulangi langkah dua sampai dengan empat sehingga sudah tidak ada lagi data yang berpindah ke kluster yang lain [Sadewo et al., 2017].



Gambar 2: Pemetaan Iterasi 1 Pada K-means



Gambar 3: Grafik Iterasi 1 Pada K-means

3. HASIL DAN PEMBAHASAN

3.1 Contoh Perhitungan K-Means

Pada bagian ini akan dijelaskan mengenai penggunaan algoritma K-Means dalam membentuk clustering (kelompok). Hal ini bertujuan membuktikan bahwa algoritma k-means mampu memberikan informasi yang dibutuhkan. Sample data sebanyak 8 record dari laporan jumlah kunjungan wisatawan ke objek wisata unggulan di Prov. DKI Jakarta. Berikut ini jumlah kunjungan wisatawan ke objek wisata unggulan pada Prov. DKI Jakarta (2007-2013) pada tabel 1:

Tabel 1. Jumlah Kunjungan Wisatawan ke Objek Wisata Unggulan

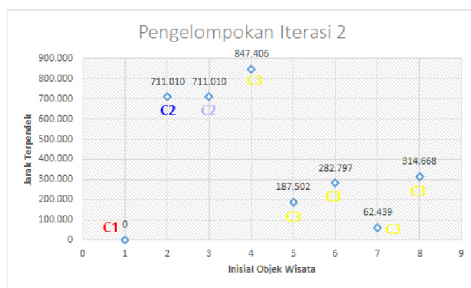
| Objek Wisata | Tahun | | | | | | |
|----------------------------|------------|------------|------------|------------|------------|------------|------------|
| | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 |
| Taman Impian Jaya Ancol | 13.377.011 | 13.567.630 | 12.920.733 | 12.834.890 | 18.450.016 | 15.848.956 | 15.948.829 |
| Taman Mini Indonesia Indah | 3.808.176 | 4.510.679 | 4.822.945 | 5.298.719 | 5.186.445 | 7.888.787 | 4.483.847 |
| Kebon Binatang Ragunan | 3.392.223 | 3.319.186 | 3.545.212 | 3.580.024 | 4.090.567 | 4.283.895 | 3.681.968 |
| Monumen Nasional | 708.757 | 924.445 | 2.112.217 | 1.253.266 | 1.516.153 | 1.418.469 | 1.380.868 |
| Museum Nasional | 157.905 | 104.739 | 165.907 | 375.710 | 193.864 | 148.118 | 169.527 |
| Museum Satria Mandala | 48.591 | 77.525 | 53.769 | 63.797 | 74.742 | 50.818 | 46.002 |
| Museum Sejarah Jakarta | 75.067 | 119.641 | 245.682 | 724.082 | 437.040 | 396.253 | 371.467 |
| Pelabuhan Sunda Kelapa | 17.217 | 14.648 | 12.677 | 34.112 | 34.179 | 32.067 | 40.210 |

Untuk membuat cluster pada tabel 1, maka terlebih dahulu ditentukan adalah sebagai berikut:

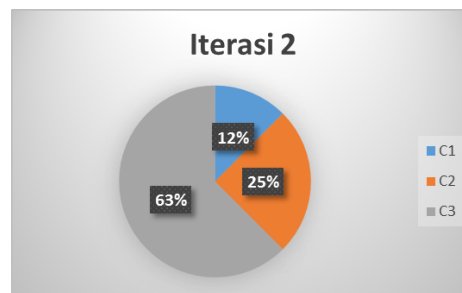
1. Menentukan jumlah cluster, jumlah cluster (kelompok) adalah 3.
2. Menentukan pusat cluster secara acak. Pada contoh kasus ini nilai pusat cluster 1 (C1)= 15.438.488, nilai pusat cluster 2 (C2) = 3.259.913, dan nilai pusat cluster 3 (C3) = 27.664.
3. Menentukan nilai cluster dari tiap data. Dalam hal ini harus menentukan nilai cluster mana yang paling dekat dengan data, maka perlu dihitung jarak setiap data dengan titik pusat setiap cluster. Pada tahap ini Distance Space digunakan untuk menghitung jarak antara data dan centroid. Adapun persamaan yang dapat digunakan salah satunya yaitu Euclidean Distance Space. Euclidean distance space digunakan dalam perhitungan jarak, hal ini dikarenakan hasil yang dapat diperoleh merupakan jarak terpendek antara dua titik yang diperhitungkan. Rumus yang digunakan adalah (1) Dari hasil perhitungan di atas maka dihasilkan nilai cluster 1, cluster 2 dan cluster ke 3 pada iterasi 1. Hasil perhitungan dari semua perhitungan iterasi 1 dapat dilihat pada Tabel 2. Proses K-means akan terus beriterasi sampai pengelompokan data

Tabel 2. Tabel Perhitungan Iterasi ke 1

| Objek Wisata | Rata-Rata | Cluster | | | Jarak | Hasil |
|----------------------------|------------|------------|------------|------------|-----------|-------|
| | | C1 | C2 | C3 | | |
| Taman Impian Jaya Ancol | 15.438.488 | 0 | 12.178.575 | 15.410.824 | 0 | C1 |
| Taman Mini Indonesia Indah | 5.175.587 | 10.262.902 | 1.915.674 | 5.147.923 | 1.915.674 | C2 |
| Kebon Binatang Ragunan | 3.753.568 | 11.684.921 | 493.655 | 3.725.904 | 493.655 | C2 |
| Monumen Nasional | 1.189.738 | 14.248.750 | 2.070.175 | 1.162.074 | 1.162.074 | C3 |
| Museum Nasional | 154.831 | 15.283.658 | 3.105.083 | 127.166 | 127.166 | C3 |
| Museum Satria Mandala | 59.536 | 15.378.953 | 3.200.378 | 31.871 | 31.871 | C3 |
| Museum Sejarah Jakarta | 279.894 | 15.158.595 | 2.980.020 | 252.229 | 252.229 | C3 |
| Pelabuhan Sunda Kelapa | 27.664 | 15.410.824 | 3.232.249 | 0 | 0 | C3 |



Gambar 4 : Pemetaan Iterasi 2 Pada K-means



Gambar 5 : Grafik Iterasi 2 Pada K-means

sama dengan pengelompokan data iterasi sebelumnya. Dengan kata lain, proses akan terus melakukan iterasi sampai data pada iterasi terakhir sama dengan iterasi sebelumnya. Pada iterasi 1 diperoleh cluster data nilai input yang dapat dilihat pada tabel 2. Pada iterasi 2, akan dilakukan pencarian nilai centroid untuk iterasi 2 dengan hasil pengelompokan pada iterasi 1. Hasil dapat diketahui pada Tabel 3: Setelah mendapatkan nilai titik tengah atau centroid,

Tabel 3. Centroid Data iterasi 2

| Atribut | Cluster 1 (C1) | Cluster 2 (C2) | Cluster 3 (C3) |
|------------------|----------------|----------------|----------------|
| Jumlah Kunjungan | 15.438.488 | 4.464.577 | 342.332 |

proses sama dilakukan dengan mencari jarak terdekat. Proses pencarian jarak terpendek, pengelompokan data pada iterasi 2 dan Clustering data dapat dilihat pada tabel 4 dan gambar (4),(5):

3.2 Analisa Data

Pada iterasi 2, pengelompokan data yang dilakukan terhadap 3 cluster dengan iterasi 1 didapatkan hasil yang sama. Dari 8 data nilai input dapat diketahui, C1 terdiri dari 1 objek wisata unggulan yakni taman impian jaya ancol, C2 terdiri dari 2 objek wisata unggulan yakni taman mini indonesia indah dan kebon binatang ragunan, C3 terdiri dari 5 objek wisata unggulan yakni monumen nasional, museum nasional, museum satria mandala, museum sejarah jakarta dan pelabuhan sunda kelapa.

4. KESIMPULAN

Untuk melakukan penilaian terhadap pengelompokan jumlah wisatawan ke objek wisata unggulan di Prov. DKI Jakarta dapat diterapkan dengan metode clustering K-Means. Data diolah untuk memperoleh jumlah wisatawan yang berkunjung ke objek wisata unggulan di Prov. DKI Jakarta. Data tersebut diolah menggunakan Aplikasi XLSTAT. Variable yang digunakan jumlah pengunjung. Data diolah dengan melakukan K-Means yang dikelompokkan menjadi 3 cluster yaitu cluster tinggi (C1) yakni jumlah kunjungan wisatawan tinggi, cluster sedang (C2) yakni

Tabel 4. Tabel perhitungan iterasi ke 2

| Objek Wisata | Rata-Rata | Cluster | | | | Hasil |
|----------------------------|------------|------------|------------|------------|---------|-------|
| | | C1 | C2 | C3 | Jarak | |
| Taman Impian Jaya Ancol | 15.438.488 | 0 | 10.973.911 | 15.096.156 | 0 | C1 |
| Taman Mini Indonesia Indah | 5.175.587 | 10.262.902 | 711.010 | 4.833.254 | 711.010 | C2 |
| Kebon Binatang Ragunan | 3.753.568 | 11.684.921 | 711.010 | 3.411.235 | 711.010 | C2 |
| Monumen Nasional | 1.189.738 | 14.248.750 | 3.274.839 | 847.406 | 847.406 | C3 |
| Museum Nasional | 154.831 | 15.283.658 | 4.309.747 | 187.502 | 187.502 | C3 |
| Museum Satria Mandala | 59.536 | 15.378.953 | 4.405.042 | 282.797 | 282.797 | C3 |
| Museum Sejarah Jakarta | 279.894 | 15.158.595 | 4.184.684 | 62.439 | 62.439 | C3 |
| Pelabuhan Sunda Kelapa | 27.664 | 15.410.824 | 4.436.913 | 314.668 | 314.668 | C3 |

jumlah kunjungan wisatawan sedang dan cluster rendah (C3) yakni jumlah kunjungan wisatawan rendah. Centroid data C1= 15.438.488, Centroid data C2= 4.464.577 dan Centroid data C3= 342.332.

Sehingga diperoleh hasil dari K-Means bahwa C1 terdiri dari 1 objek wisata unggulan yakni Taman Impian Jaya Ancol, C2 terdiri dari 2 objek wisata unggulan yakni Taman Mini Indonesia Indah Dan Kebon Binatang Ragunan, C3 terdiri dari 5 objek wisata unggulan yakni Monumen Nasional, Museum Nasional, Museum Satria Mandala, Museum Sejarah Jakarta Dan Pelabuhan Sunda Kelapa. Hasil dari penelitian menunjukkan bahwa 5 objek wisata unggulan yang berada di cluster paling rendah menjadi catatan bagi pemerintah Prov. DKI Jakarta. Dengan adanya perbaikan sarana dan prasarana objek wisatawan unggulan dapat meningkatkan jumlah kunjungan wisatawan yang berdampak pada pengenalan objek wisata dan peningkatan devisa negara.

DAFTAR PUSTAKA

References

- Afrisawati. Implementasi data mining pemilihan pelanggan potensial menggunakan algoritma k-means. *Jurnal Pelita Informatika Budi Karma*, 5(3), 2013.
- Nielza Atthina and Lizda Iswari. Klasterisasi data kesehatan penduduk untuk menentukan rentang derajat kesehatan daerah dengan metode k-means. In *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, volume 1, 2014.
- Emha Taufiq Luthfi Kusri. *Algoritma data mining*. Yogyakarta: Andi Offset, 2009.
- Ahlihi Masruro, Kusri Kusri, and Emha Taufiq Luthfi. Sistem penunjang keputusan penentuan lokasi wisata menggunakan k-means clustering dan topsis. *Data Manajemen dan Teknologi Informasi*, 15(4), 2014.
- Johan Oscar Ong. Implementasi algoritma k-means clustering untuk menentukan strategi marketing president university. *Jurnal Ilmiah Teknik Industri*, 12(1):10–20, 2013.
- Eko Prasetyo. *Data mining konsep dan aplikasi menggunakan matlab*. Yogyakarta: Andi, 2012.
- Mhd Gading Sadewo, Agus Perdana Windarto, and Dedy Hartama. Penerapan datamining pada populasi daging ayam ras pedaging di indonesia berdasarkan provinsi menggunakan k-means clustering. *InfoTekJar (Jurnal Nasional Informatika dan Teknologi Jaringan)*, 2(1): 60–67, 2017.
- Neha Soni and Amit Ganatra. Categorization of several clustering algorithms from different perspective: a review. *International Journal of*, 2012.
- Suprihatin. Klastering k-means untuk penentuan nilai ujian. *JUSI*, 1(1), 2011.

Bambang Tri Wahyu and Angga Anggriawan. Sistem rekomendasi paket wisata se-malang raya menggunakan metode hybrid content based and collaborative. *Jurnal Ilmiah Teknologi Informasi Asia*, 9(1):6–13, 2015.

Anindya Khrisna Wardhani. K-means algorithm implementation for clustering of patients disease in kajen clinic of pekalongan. *Jurnal Transformatika*, 14(1):30–37, 2016.

Agus Perdana Windarto. Implementation of data mining on rice imports by major country of origin using algorithm using k-means clustering method. *International Journal of Artificial Intelligence Research*, 1(2), 2017.
