

Penerapan Metode K-Means untuk Pemetaan Objek Wisata Sebagai Rekomendasi Prioritas Pengembangan Pariwisata

Dwi Rahmiyati¹⁾, Nuraeni²⁾, Ervan Britantono Siswantoro³⁾

^{1,2)} Universitas Binaniaga Indonesia / Sistem Informasi

³⁾ Universitas Binaniaga Indonesia / Teknologi Informasi

¹⁾ dwi.rahmiyati@gmail.com[✉], ²⁾ eni.nuraeni3005@gmail.com,

³⁾ ervan.britantono@gmail.com

ABSTRACT

Tourism is one of the driving factors of economic growth in Indonesia. Uneven destination development and the absence of clear mapping can have an impact on inappropriate allocation of funds and development priorities. Mapping tourist attractions can help related parties in determining tourist attractions that are priority for development. The purpose of this study was to obtain accurate mapping of tourist attractions and develop a prototype application for implementing K-Means to help determine recommendations for priority tourism development. The source of research data used was tourist attraction data in the Bogor area. The variables used include Number of Visits, Facilities, Public Vehicle Accessibility, Entrance Ticket Prices, and Area. The results of the K-Means calculation are cluster 1 consisting of 4 tourist attractions, cluster 2 consisting of 10 and cluster 3 consisting of 81 tourist attractions. Based on the feasibility test on the application that was built, the results were 100% or very feasible. User testing with the PSSUQ questionnaire obtained an Overall value of 86% which means it is very feasible to use. The validity test of the cluster using the silhouette coefficient obtained a result of 0.78 in the Silhouette Category table, it was concluded that the cluster created was included in the "Strong Structure" category.

Keywords: K-Means Algorithm, Mapping, Silhouette Coefficient, Tourist Attractions, Tourism Development.

ABSTRAK

Pariwisata menjadi salah satu faktor penggerak pertumbuhan ekonomi di Indonesia. Perkembangan destinasi yang tidak merata serta tidak adanya pemetaan yang jelas dapat berdampak pada penyalokan dana dan prioritas pengembangan yang kurang tepat. Pemetaan objek wisata dapat membantu pihak terkait dalam upaya menentukan objek wisata yang menjadi prioritas pengembangan. Tujuan penelitian dimaksudkan untuk mendapatkan pemetaan objek wisata yang akurat dan mengembangkan prototype aplikasi penerapan K-Means guna membantu menentukan rekomendasi prioritas pengembangan wisata. Sumber data penelitian yang digunakan adalah data objek wisata di wilayah Bogor. Variabel yang digunakan meliputi Jumlah Kunjungan, Fasilitas, Aksesibilitas Kendaraan Umum, Harga Tiket Masuk, dan Luas Wilayah. Hasil perhitungan K-Means yaitu cluster 1 berjumlah 4 objek wisata, cluster 2 berjumlah 10 dan cluster 3 berjumlah 81 objek wisata. Berdasarkan uji kelayakan pada aplikasi yang dibangun diperoleh hasil 100% atau sangat layak. Uji pengguna dengan kuisioner PSSUQ diperoleh nilai Overall 86% yang berarti sangat layak digunakan. Uji validitas cluster menggunakan *silhouette coefficient* di dapat hasil 0.78 dalam tabel Kategori *Silhouette* disimpulkan cluster yang dibuat termasuk dalam kategori "Strong Structure".

Kata kunci: Algoritma K-Means, Pemetaan, *Silhouette Coefficient*, Objek Wisata, Pengembangan Pariwisata.

I. PENDAHULUAN

Pariwisata memiliki potensi besar baik keindahan alam maupun keragaman budayanya. Hal ini memberikan kontribusi terhadap pembangunan ekonomi serta peningkatan pendapatan nasional dan

masyarakat lokal. Berdasarkan UU No. 10 Tahun 2009 tentang Kepariwisata, dijelaskan bahwa pariwisata adalah macam-macam kegiatan wisata dan didukung berbagai fasilitas serta layanan yang disediakan oleh masyarakat, pengusaha, pemerintah, dan pemerintah daerah. Dinas Pariwisata dan Kebudayaan memiliki

program kegiatan rencana kerja pengembangan daya tarik wisata. Tujuannya untuk mengembangkan objek wisata dengan harapan dapat meningkatkan jumlah kunjungan wisatawan nusantara maupun mancanegara. Namun, pengembangan sektor ini juga menghadapi tantangan, seperti kebutuhan untuk meningkatkan kualitas objek wisata dan mengembangkan inovasi dalam pengelolaan pariwisata. Perkembangan destinasi pariwisata nasional di Indonesia belum sepenuhnya merata [1]. Tantangan juga dihadapi di Kabupaten Bogor, yaitu potensi daya tarik wisata belum dikelola secara optimal, perlunya peningkatan sarana dan prasarana pariwisata serta perbaikan akses menuju objek wisata [2].

Data mining adalah proses pengolahan data skala besar [3]. Proses ini mencari hubungan antar data pada pola yang sering terjadi maupun jarang terjadi. Tujuannya untuk mengekstrak informasi dari big data sehingga dihasilkan pola informasi yang menarik sebagai pengetahuan baru yang tersembunyi [4]. *Clustering* yaitu pengelompokan suatu set objek yang mirip kedalam kelas atau kelompok data dikelompokkan kedalam kelas atau kelas yang memiliki kemiripan tinggi satu sama lain tetapi memiliki perbedaan pada *cluster* lainnya [5]. Dengan analisa *cluster* akan lebih focus dalam Analisa lebih lanjut [6].

Marini dan Suhendra melakukan penelitian menggunakan algoritma K-Means untuk pemetaan *cluster* daerah pariwisata di Kabupaten Teluk Wondama [7]. Proses penentuan *cluster* ideal menggunakan metode *Elbow* dan *Silhouette*. Banyak potensi wisata kurang dioptimalkan karena jarak yang jauh di setiap distrik. Hasil penelitian didapatkan tiga distrik pada *cluster* pertama dan lima distrik pada *cluster* dua. Penelitian tersebut mengelompokkan distrik-distrik yang potensial untuk dimaksimalkan oleh pemerintah daerah.

Penelitian lain dilakukan oleh Al-Fahmi pada pariwisata Bojonegoro dengan menggunakan metode K-Means [8]. *Data understanding* yang digunakan adalah jumlah kunjungan, harga tiket masuk, pertumbuhan jumlah kunjungan, dan luas wilayah. Dihasilkan 3 *cluster* yaitu maju, berkembang dan rintisan. Disimpulkan bahwa metode K-Means dapat membantu pengambilan keputusan strategi pemasaran yang sesuai. Penelitian Purnomo dan Prasetyaningrum menggunakan metode K-Means untuk mengklasifikasi kunjungan wisatawan di Kota Yogyakarta dengan *cluster* kunjungan rendah, sedang dan tinggi. Penelitian tersebut diharapkan membantu Dinas Pariwisata mengembangkan potensi pariwisata untuk meningkatkan minat kunjungan wisatawan [9].

Berdasarkan hal tersebut penelitian untuk memetakan objek wisata sangat penting dilakukan. *Data understanding* yang digunakan meliputi jumlah kunjungan, fasilitas, aksesibilitas kendaraan umum, harga tiket masuk dan luas wilayah. Penelitian ini merekomendasikan pemetaan dengan metode K-Means sebagai alat penting untuk mengidentifikasi potensi wisata yang ada, menentukan prioritas pengembangan,

serta merencanakan strategi yang lebih efektif dalam pengelolaan dan promosi pariwisata.

Sistem informasi dikembangkan menggunakan metode *System Development Life Cycle (SDLC)*. Aplikasi yang dikembangkan berbasis web dengan bahasa pemrograman *python* dengan *library numpy* dan *framework web Laravel*. *Python* memiliki modul siap pakai [10]. *Framework Laravel* merupakan keluaran dari PHP berlisensi MIT [11]. Manfaatnya memberi kemudahan dalam proses *maintenance* dan *scalability* [12]. Metode K-Means diharapkan mampu mendukung proses pemetaan lebih tepat. Berdasarkan system yang dibuat dilakukan pengukuran oleh ahli sitem dengan *whitebox*, uji kelayakan pengguna dengan PSSUQ. Hasil persentase akan memberikan jawaban terkait kelayakan dari aspek yang diteliti [13]. Uji validitas *cluster* yang dilakukan menggunakan *silhouette coefficient*.

II. METODE

A. Algoritma K-Means

Metode penelitian digunakan untuk memecahkan dan merekomendasikan masalah. K-Means adalah teknik pengelompokan berbasis jarak dengan membagi data menjadi beberapa klaster [14]. Karakteristik yang dimiliki cenderung sama setiap kelompok. Data harus digali menjadi informasi baru yang dapat membantu dalam pengambilan Keputusan [15]. *Cluster* optimal inilah yang akan digunakan sebagai penentuan strategi bisnis [16]. Dengan pemetaan yang tepat, diharapkan penelitian dengan K-Means ini dapat memberi kontribusi dalam mengetahui lokasi-lokasi potensial yang memerlukan pengembangan, mengelola aliran wisatawan secara lebih efisien, dan melestarikan sumber daya alam dan budaya secara berkelanjutan.

Pemrosesan diawali dengan pengelompokan *centroid* pertama sebagai titik awal untuk setiap *cluster* yang dipilih secara acak, kemudian dihitung ulang hingga tidak ada perubahan kelompok *cluster*. Langkah-langkah algoritma K-Means sebagai berikut:

- 1) Tentukan jumlah *cluster* (k) pada data set;
- 2) Tentukan nilai pusat (*centroid*); penentuan nilai *centroid* pada tahap awal dilakukan secara random, sedangkan pada tahap iterasi digunakan rumus seperti pada persamaan berikut ini:

$$V_{ij} = \frac{1}{n_i} \sum_{k=0}^{n_i} x_{kj}$$

dengan V_{ij} adalah *centroid* rata-rata *cluster* ke- i untuk variabel ke- j , N_i adalah jumlah anggota *cluster* ke- i , K adalah indeks dari *cluster*, J adalah indeks dari variable, dan X_{kj} adalah nilai data ke- k variabel ke- j untuk *cluster* tersebut.

- 3) Pada masing-masing record, hitung jarak terdekat dengan *centroid* jarak *centroid* yang digunakan adalah *euclidean distance*, dengan rumus pada persamaan dibawah ini:

$$De = \sqrt{(xi - si)^2 + (yi - ti)^2}$$

dengan De adalah euclidean distance, I adalah banyaknya objek, (x, y) adalah koordinat objek, dan (s, t) adalah koordinat *centroid*.

- 4) Kelompokkan objek berdasarkan jarak ke *centroid* terdekat, ulangi langkah ke-3 hingga langkah ke-4, lakukan iterasi hingga *centroid* bernilai optimal.

B. Teknik Analisis Data

Data yang digunakan adalah data objek wisata di Bogor bersumber dari Dinas Pariwisata dan Kebudayaan. Sebagai pendukung data sekunder didapat dari berbagai portal data pemerintah yang tersedia secara umum digunakan sesuai relevansinya. Metode pengujian yang diterapkan dalam penelitian ini adalah metode *silhouette coefficient*. Metode ini akan menguji kualitas dari setiap *cluster* yang dihasilkan dengan menggabungkan metode *cohesion* dan *separation*. *Silhouette Coefficient* membantu melihat kualitas dan kekuatan suatu *cluster* [17].

Langkah yang perlu dilakukan untuk menghitung *Silhouette Coefficient* [18], yaitu:

- 1) Untuk setiap objek I, hitung rata-rata jarak dari objek I dengan seluruh objek yang berada dalam satu cluster disebut dengan a
- 2) Untuk setiap objek I, dihitung rata – rata jari dari objek I dengan objek yang berada di cluster lainnya. Dari semua jarak rata– rata tersebut diambil nilai yang paling kecil, nilai ini disebut dengan bj (kuadrat kecil)
- 3) Setelah itu maka nilai *Silhouette Coefficient* dari objek i adalah:

$$S_i = (b_i - a_i) / \max(a_i, b_i)$$

dengan a(i) adalah rata – rata jarak objek i terhadap seluruh objek di dalam *cluster* dan b(i) adalah rata – rata jarak objek terhadap seluruh objek di luar *cluster*.

Berikut tabel kategori *Silhouette Coefficient* menurut Kauffman dan Rousseeuw [19].

Tabel 1. *Silhouette Coefficient*

Nilai <i>Silhouette Coefficient</i>	Keterangan
0,71 - 1	Strong Structure
0,51 - 0,70	Good Structure
0,26 - 0,50	Weak Structure
<= 0,25	No Structure

III. HASIL DAN PEMBAHASAN

A. Hasil

Metode yang digunakan adalah metode klasterisasi Algoritma K-Means yang akan digunakan untuk mengelompokan data dan menghasilkan klistor. Tabel 2 merupakan data objek wisata di Bogor.

Tabel 2. Data Objek Wisata

Kode	Nama Objek wisata	Jumlah kunjungan (Orang)	Fasilitas	Aksesibilitas Kendaraan umum	Harga Tiket Masuk (Rupiah)	Luas Wilayah (Hektar)
1	Curug Kembar batu layang	599	Cukup	Ada	10000	10
2	Buper Citamiang	1043	Cukup	Ada	40000	10
3	Curug Cipamingkis	38996	Baik	Tidak Ada	35000	16.5
4	Curug Ciherang	42207	Cukup	Tidak ada	40000	73
5	Curug Cibereum	1034	Cukup	Ada	18500	0.25
6	Curug cIdulang	2143	Cukup	Ada	15000	35
7	Curug Barong Leuwi Hejo	4043	Minim	Tidak Ada	25000	35
8	Track sepeda puncak kondang	538	Cukup	Ada	10000	256.7
9	Curug Putri Kencana	6905	Cukup	Tidak Ada	25000	5
10	WA baru Jeruk	1007	Cukup	Ada	15000	1
:	:	:	:	:	:	:
95	Kuntum Farmfield	21704	Cukup	Ada	60000	10

Variabel yang digunakan dalam pemetaan objek wisata untuk rekomendasi prioritas pengembangan pariwisata adalah Jumlah Kunjungan, Fasilitas, Aksesibilitas Kendaraan Umum, Harga Tiket Masuk, dan Luas Wilayah. Secara umum algoritma K-Means bertujuan untuk membagi data menjadi beberapa kelompok, dengan tahapan sebagai berikut:

- 1) Menyiapkan Dataset

Dataset terdiri dari 95 baris dan 6 kolom. Variabel yang di tranformasi yaitu variabel fasilitas dan Aksesibilitas kendaraan umum.

Tabel 3. Transformasi Variabel Fasilitas

Variabel Fasilitas	
Data Awal	Data Transformasi
Minim	1
Cukup	2
Baik	3
Baik	3

Fasilitas kategori minim yaitu objek wisata dengan fasilitas kurang dari cukup. Fasilitas kategori cukup yaitu tersedia fasilitas area parkir, toilet, mushola dan warung. Sedangkan untuk fasilitas baik meliputi kategori cukup dengan tambahan fasilitas seperti petugas medis, petugas keamanan ataupun pusat informasi.

Tabel 4. Transformasi Variabel Aksesibilitas

Variabel Aksesibilitas Kendaraan Umum	
Data Awal	Data Transformasi
Tidak Ada	1
Ada	2

Tabel 4 menunjukkan objek wisata yang memiliki akses transportasi kendaraan umum dan yang tidak tersedia.

2) Menentukan Jumlah Kluster

Dari dataset objek wisata di Bogor akan dibentuk 3 kluster yaitu kluster prioritas rendah, kluster prioritas sedang dan kluster prioritas tinggi.

3) Menentukan Titik Centroid

Data yang digunakan pada *centroid* 1 adalah data baris ke 47, *centroid* 2 adalah data pada baris ke 32 dan untuk *centroid* 3 yaitu data pada baris ke 24.

Tabel 5. Titik Awal Centroid

Centroid	Jumlah Kunjungan	Fasilitas	Aksesibilitas Kendaraan umum	HTM	Luas Wilayah
C1	15881	2	2	300000	6
C2	1024	2	2	15000	22
C3	257	2	2	10000	1,5

4) Hitung jarak Centroid

Perhitungan Jarak dengan *Centroid* 1

Data ke 1 :

$$D(X_1, C_1) = \sqrt{(599 - 15881)^2 + (2 - 2)^2 + (2 - 2)^2 + (10000 - 300000)^2 + (10 - 6)^2}$$

$$(X_1, C_1) = 84333539540$$

Data ke 2 :

$$D(X_2, C_1) = \sqrt{(1043 - 15881)^2 + (2 - 2)^2 + (2 - 2)^2 + (40000 - 300000)^2 + (10 - 6)^2}$$

$$= 67820166260$$

....

Data ke 95 :

$$D(X_{95}, C_1) = \sqrt{(21704 - 15881)^2 + (2 - 2)^2 + (2 - 2)^2 + (60000 - 300000)^2 + (10 - 6)^2}$$

$$= 57633907345$$

Perhitungan Jarak dengan *Centroid* 2

Data ke 1 :

$$D(X_1, C_2) = \sqrt{(599 - 1024)^2 + (2 - 2)^2 + (2 - 2)^2 + (10000 - 15000)^2 + (10 - 22)^2}$$

$$D = 25180769$$

Data ke 2 :

$$D(X_2, C_2) = \sqrt{(1043 - 1024)^2 + (2 - 2)^2 + (2 - 2)^2 + (40000 - 15000)^2 + (10 - 22)^2}$$

$$D(X_1, C_1) = 625000505$$

...

Data ke 95 :

$$D(X_{95}, C_2) = \sqrt{(21704 - 1024)^2 + (2 - 2)^2 + (2 - 2)^2 + (60000 - 15000)^2 + (10 - 22)^2}$$

$$(X_{95}, C_1) = 2452662544$$

Perhitungan Jarak dengan *Centroid* 3

Data ke 1 :

$$D(X_1, C_3) = \sqrt{(599 - 257)^2 + (2 - 2)^2 + (2 - 2)^2 + (10000 - 10000)^2 + (10 - 1.5)^2}$$

$$D(X_1, C_1) = 117036$$

Data ke 2 :

$$D(X_2, C_3) = \sqrt{(1043 - 257)^2 + (2 - 2)^2 + (2 - 2)^2 + (40000 - 10000)^2 + (10 - 1.5)^2}$$

$$D(X_1, C_1) = 900617868$$

...

Data ke 95 :

$$D(X_{95}, C_3) = \sqrt{(21704 - 257)^2 + (2 - 2)^2 + (2 - 2)^2 + (60000 - 10000)^2 + (10 - 1.5)^2}$$

$$D(X_1, C_1) = 2959973881$$

Perhitungan di atas seharusnya dihitung keseluruhan data. Berikut hasil dari perhitungan *centroid* 1, *centroid* 2 dan *centroid* 3, jarak terdekat dan klasternya.

Tabel 6. Tabel Hasil Perhitungan *Centroid* iterasi 1

Kode Objek Wisata	C1	C2	C3	Jarak Terdekat	Cluster
1	84333539540	25180769	117036	117036	C3
2	67820166260	625000505	900617868	625000505	C2
3	70759303337	1841872816	2125710348	1841872816	C2
4	68293062766	2321042091	2659807613	2321042091	C2

Kode Objek Wisata	C1	C2	C3	Jarak Terdekat	Cluster
5	7946268 3442	122505 73	728537 31	122505 73	C2
6	8141373 3485	125233 0	285581 18	125233 0	C2
7	7576513 9087	109114 532	239334 920	109114 532	C2
8	8433547 0499	252912 80	144088	144088	C3
9	7570556 8578	134586 451	269195 917	134586 451	C2
10	8144623 5901	730	255625 00	730	C2
:	:	:	:	:	:
95	5763390 7345	245266 2544	295997 381	245266 2544	C2

Tabel 6 menunjukkan hasil perhitungan dengan rumus *Euclidean Distance* antara tiap variabel dengan *centroid Cluster 1 (C1), Cluster 2 (C2)* dan *Cluster 3 (C3)* sehingga didapat jarak terdekat dari masing-masing atribut (data N), kemudian untuk ditentukan masuk ke *cluster* masing – masing.

5) Menentukan Posisi Cluster

Setelah proses pengelompokan data terdapat 5 objek wisata pada *cluster 1*, terdapat 70 objek wisata pada *cluster 2* dan 20 objek wisata pada *cluster 3*. Selanjutnya dilakukan penentuan titik *centroid* baru, dengan menggunakan rumus :

Mencari nilai *centroid* klaster 1 baru :

$$\begin{aligned} \mu_1 &= \left(\frac{1691625}{5}\right) & \mu_3 &= \left(\frac{10}{5}\right) \\ &= 338325 & &= 2 \\ \mu_2 &= \left(\frac{10}{5}\right) & \mu_4 &= \left(\frac{1369000}{5}\right) \\ &= 2 & &= 273800 \\ \mu_5 &= \left(\frac{168}{5}\right) \\ &= 33,6 \end{aligned}$$

Mencari nilai *centroid* klaster 2 baru :

$$\begin{aligned} \mu_1 &= \left(\frac{6238894}{70}\right) & \mu_4 &= \left(\frac{1627000}{70}\right) \\ &= 338325 & &= 23242,85714 \\ \mu_2 &= \left(\frac{141}{70}\right) & \mu_5 &= \left(\frac{31588,407}{70}\right) \\ &= 2,014285714 & &= 451,2629571 \\ \mu_3 &= \left(\frac{122}{70}\right) \\ &= 1,742857143 \end{aligned}$$

Mencari nilai *centroid* klaster 3 baru :

$$\begin{aligned} \mu_1 &= \left(\frac{128726}{20}\right) & \mu_4 &= \left(\frac{160000}{20}\right) \\ &= 6436,3 & &= 8000 \\ \mu_2 &= \left(\frac{40}{20}\right) & \mu_5 &= \left(\frac{1335}{20}\right) \\ &= 2 & &= 66,7492 \\ \mu_3 &= \left(\frac{34}{20}\right) \\ &= 1,7 \end{aligned}$$

Centroid baru ini merupakan hasil dari menghitung rata-rata nilai dari semua anggota pada klaster. Mengulang perhitungan jarak data dengan *centroid* hingga tidak ada lagi perubahan jumlah data pada setiap anggota *cluster*. Dari perhitungan yang dilakukan pada penelitian ini terjadi iterasi sebanyak 6 kali, data pertama hingga data yang terakhir tidak ada lagi yang bergeser dari satu *cluster* ke *cluster* lain.

Tabel 7. Titik Centroid Iterasi Keenam

Centroid	Jumlah Kunjungan	Fasilitas	Aksesibilitas Kendaraan umum	HTM	Luas Wilayah
1	1113617. 25	2	2	61375	674.12 5
2	208477.9	2.1	2	1458 50	272.39
3	18765.39 506	2	1.7037037 04	1792 5.925 93	341.61 72

Hasil akhir perhitungan di dapatkan jumlah data setiap klasternya adalah untuk *cluster 1* dengan prioritas rendah berjumlah 4 objek wisata yaitu Argowisata Gunung Mas, Cimory Dairy Land (Dairyland Puncak), Taman Safari Indonesia dan Kebun Raya. *Cluster 2* dengan prioritas sedang berjumlah 10 objek wisata terdiri dari Alamanda Indonesia/ Bogor Rafting, Camp Hulu Cai, Cibalong Happy Land, Cimory Riverside (Diary Land Riverside), Inagro, Jungleland, Nicole's River Park, Taman Wisata Matahari, The Ciliwung tea estate, The Jungledan. Sedangkan *cluster 3* berjumlah 81 objek wisata. Dari hasil tersebut maka yang menjadi rekomendasi prioritas pengembangan yaitu pada *cluster 3* dengan prioritas tinggi. Penggunaan variabel meliputi jumlah kunjungan, fasilitas, aksesibilitas kendaraan umum, harga tiket masuk, dan luas wilayah diharapkan akan menghasilkan pemetaan yang lebih terperinci untuk membantu menganalisa guna pengambilan keputusan pengembangan yang lebih efisien.

B. Uji Hasil

Berdasarkan hasil pengembangan terkait metode yang diterapkan pada sistem informasi. Hasil klaster menggunakan metode k-means diuji dengan menggunakan metode *silhouette coefficient*. Metode ini bertujuan untuk menguji kualitas *cluster*.

- 1) Menghitung jarak rata – rata dari suatu data dengan semua data pada klaster yang sama. Misalnya data pertama.

- a. Data Pertama terhadap seluruh data yang termasuk dalam kluster yang sama yaitu kluster 1:

$$a(1)_1 = \sqrt{\frac{(793171 - 1061668)^2 + (2 - 2)^2 + (2 - 2)^2 + (30000 - 15500)^2 + (3.5 - 2551)^2 + \dots + (1011974 - 1061668)^2 + (2 - 2)^2 + (2 - 2)^2 + (30000 - 15500)^2 + (87 - 2551)^2}{4}}$$

$$a(1)_1 = \sqrt{\frac{876068.9045}{4}}$$

$$a(1)_1 = 219017.2261$$

- b. Data Pertama terhadap seluruh data yang termasuk dalam kluster yang sama yaitu kluster 2:

$$a(1)_2 = \sqrt{\frac{(3812 - 1000)^2 + (2 - 2)^2 + (2 - 2)^2 + (300000 - 269000)^2 + (25 - 2)^2 + \dots + (471889 - 1000)^2 + (2 - 2)^2 + (2 - 2)^2 + (50000 - 269000)^2 + (4 - 2)^2}{10}}$$

$$a(1)_2 = \sqrt{\frac{2536806.801}{10}}$$

$$a(1)_2 = 253680.6801$$

- c. Data Pertama terhadap seluruh data yang termasuk dalam kluster yang sama yaitu kluster 3:

$$a(1)_3 = \sqrt{\frac{(1043 - 599)^2 + (2 - 2)^2 + (2 - 2)^2 + (40000 - 10000)^2 + (10 - 10)^2 + \dots + (21704 - 599)^2 + (2 - 2)^2 + (2 - 2)^2 + (60000 - 10000)^2 + (10 - 10)^2}{81}}$$

$$a(1)_3 = \sqrt{\frac{2235716.061}{81}}$$

$$a(1)_3 = 27601.43285$$

- 2) Melakukan perhitungan jarak rata – rata dari suatu data terhadap semua data dari *cluster* yang berbeda. Dan di ambil jarak terkecilnya. Pada langkah ini menghitung nilai b(i).

- a. Kluster 1 ke kluster 2

$$b(1)_{1,2} = \sqrt{(1000 - 1061668)^2 + (2 - 2)^2 + (2 - 2)^2 + (269000 - 15500)^2 + (2 - 10)^2 + \dots + (471889 - 1061668)^2 + (2 - 2)^2 + (2 - 2)^2 + (50000 - 15500)^2 + (4 - 10)^2}$$

$$b(1)_{1,2} = 868957.2837$$

Perhitungan ini terus dilakukan hingga keseluruhan data pada kluster 1 dan kluster 2

- b. Kluster 1 ke kluster 3

$$b(1)_{1,3} = \sqrt{(599 - 1061668)^2 + (2 - 2)^2 + (2 - 2)^2 + (10000 - 15500)^2 + (10 - 10)^2 + \dots + (21704 - 1061668)^2 + (2 - 2)^2 + (2 - 2)^2 + (60000 - 15500)^2 + (10 - 10)^2}$$

$$b(1)_{1,3} = 1043114.455$$

Perhitungan ini terus dilakukan hingga keseluruhan data pada kluster 1 dan kluster 3.

- 3) Setelah mengetahui nilai a(i) dan b(i), maka nilai silhouette coefficient, yaitu

$$s(1) = 1 - \frac{a(i)}{b(i)}$$

$$s(1) = 1 - \frac{219017.2261}{868957.2837} = 0.747953979$$

Perhitungan nilai *silhouette coefficient* dilakukan pada semua data. Nilai *silhouette coefficient* tersebut dapat dilihat pada tabel 8.

Tabel 8. Hasil *Silhouette Coefficient*

Kode Objek Wisata	a(i)	b(i)	s(i)
1	27601.43285	316575.2774	0.912812418
2	36842.34905	302091.8679	0.878042566
3	38909.35314	281799.8882	0.861925591
4	42857.14815	277705.2687	0.845673983
5	26223.50162	312097.9481	0.915976693
6	25623.33618	313132.0021	0.918170816
7	27298.27794	313132.0021	0.91282182
8	27636.70377	316612.6945	0.912711321
9	26775.03047	305358.5099	0.912316082
10	25972.40356	313830.9452	0.917240782
:	:	:	:
95	45441.17655	265508.9058	0.625550173

- 4) Menghitung nilai rata – rata *silhouette coefficient* pada semua kluster sebagai berikut:

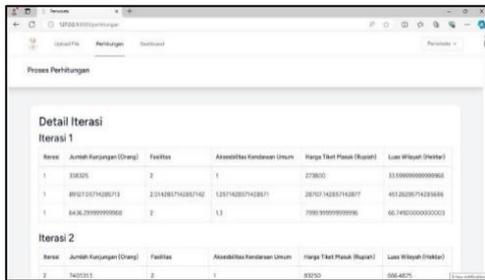
$$s(i) = \frac{0.912812418 + 0.878042566 + 0.861925591 + \dots + 0.625550173}{95}$$

$$s(i) = 0.78$$

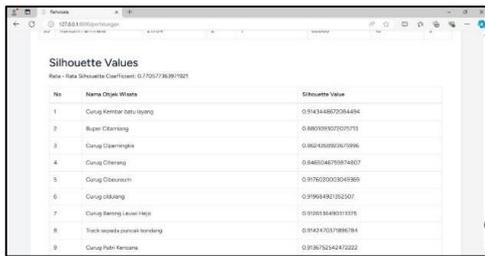
Maka diperoleh hasil *silhouette coefficient* sebesar 0.78 dengan kategori kluster kuat. Nilai positif dan mendekati 1 tingkat kesesuaian objek dengan klasternya.

C. Desain Aplikasi

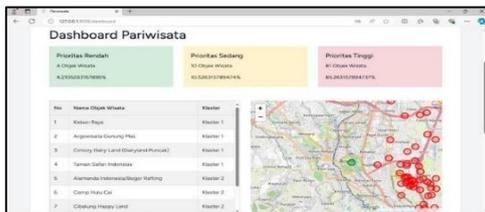
Aplikasi dikembangkan untuk memodelkan dan mengatur sistem agar mendapatkan *output* yang diharapkan. Interface terdiri dari halaman login, halaman *upload file*, halaman perhitungan dan halaman *dashboard*.



Gambar 1. Halaman perhitungan detail iterasi



Gambar 2. Halaman Perhitungan Silhouette



Gambar 3. Halama dashboard

D. Uji Coba Produk

Uji produk aplikasi kepada ahli sistem dan pengguna menggunakan persentase kelayakan, dengan rumus:

$$\text{Persentase kelayakan} = \frac{\text{skor yang diobservasi}}{\text{skor yang diharapkan}} \times 100\%$$

Hasil pengujian whitebox dari ahli sistem diperoleh 100% dikategorikan “sangat layak”. Sedangkan uji pengguna menggunakan PSSUQ 16 didapatkan hasil 85% dikategorikan “sangat layak”.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang diuraikan, maka kesimpulan yaitu menerapkan metode Algoritma K-means dapat memberikan pemetaan objek wisata untuk rekomendasi prioritas pengembangan pariwisata dengan akurat karena telah diuji oleh silhouette coefficient. Menerapkan metode Algoritma K-Means dalam pemetaan objek wisata untuk rekomendasi pengembangan pariwisata dimasa yang akan datang menjadi lebih efektif dari proses yang dilakukan sebelumnya. Hasil pengembangan prototype pada pemetaan objek wisata untuk untuk rekomendasi pengembangan pariwisata menampilkan proses perhitungan algoritma k-means, hasil kluster dan hasil pemetaan objek wisata berdasarkan tingkat prioritas pengembangan. Hasil uji akurasi dengan silhouette coefficient sebesar 0.78 yang berarti termasuk kedalam kategori kekuatan kluster kuat, kemudian hasil

kuisisioner sebesar 85%, serta hasil kuisisioner kepada ahli sebesar 100% maka dapat dikategorikan sangat layak digunakan. Terdapat beberapa saran guna meningkatkan akurasi hasil pemetaan. Disarankan dapat mempertimbangkan variabel lain yang relevan dan mengkombinasikan penggunaan metode prioritas seperti sistem penunjang keputusan untuk dapat lebih menentukan urutan prioritas per objek sehingga hasil lebih optimal.

REFERENSI

- [1] M.F. Masteriarsa and Riyanto, “Pemetaan Destinasi Pariwisata Berdasarkan Daya Dukung Kepariwisataannya Provinsi di Indonesia”, Jurnal TAMBORA (Science and Technology/Social Humaniora), Vol.7, No. 2, pp 1-9, Juli 2023.
- [2] “Dinas Kebudayaan dan Pariwisata Kabupaten Bogor”, 2021, Perubahan Rencana Strategi Dinas Kebudayaan dan Pariwisata Kabupaten Bogor Tahun 2018-2023, <https://disbudpar.bogorkab.go.id/wp-content/uploads/2024/02/Renstra-Perubahan-2018-2023.pdf>, diakses tanggal 6 Februari 2025.
- [3] F. Fahrizal, B. Irawan and A. Bahtiar, “Analisis Produk Terlaris Dan Pengujian K-Means Untuk Umkm Cetom”, JATI (Jurnal Mahasiswa Teknik Informatika), Vol. 8 No. 3, pp 3055-3061, Juni 2024.
- [4] A. Wanto, M.N Siregar, A.P Windarto, D. Hartama, N.L.W.S.R Ginantra, D. Napitupulu, E.S Negara, M.R. Lubis, S.V. Dewi and C. Prianto, Data Mining Algoritma & Implementasi, Yayasan Kita Menulis, 2020.
- [5] P. Bhatia, Data Mining and Data Warehousing, Cambridge University Press, 2019.
- [6] J. Han, M. Kamber and J. Peñ, Data Mining Concepts and Techniques, “In Data Mining Concepts and Techniques”, Waltham: Elsevier Inc, 2012.
- [7] L.F. Marini and C.D Suhendra, “Penggunaan Algoritma K-Means Pada Aplikasi Pemetaan Kluster Daerah Pariwisata”, Jurna Media Informatika Budidarma, Vol. 7, No. 2, pp 707-713, April 2023.
- [8] B.M. Al-Fahmi, E. Rahmawati and T. Sagirani, “Penerapan K-Means Clustering Pada Pariwisata Kabupaten Bojonegoro Untuk Mendukung Keputusan Strategi Pemasaran”. Jurnal Nasional Teknologi dan Sistem Informasi, Vol. 9, No. 2, pp 141-149, Agustus 2023.
- [9] B.S. Purnomo and P.T. Prasetyaningrum, “Penerapan Datamining Dalam Mengelompokkan Kunjungan Wisatawan Di Kota Yogyakarta Menggunakan Metode K-Means”, Journal of Computer Science and Technology (JCS-TECH), Vol. 1, No. 1, pp 27-32, November 2021
- [10] F.D. Adhinata and G.F. Fitriana, Pengenalan Bahasa Pemrograman Python Dan Implementasi Machine Learning, Teknosain, 2022.
- [11] F. Abdussalaam and S. A. Saputra, “Perancangan Sistem Informasi Complaint Management Dengan Metode Rad Menggunakan Framework Laravel,” J. E-Komtek, vol. 2, no. 2, pp. 54–68, 2018.
- [12] D. Wijonarko and F. W. S. Budi, “Implementasi Framework Laravel Dalam Sistem Pendaftaran,” Jurnal Inform. Rekayasa Elektron (JIRE), vol. 2, no. 2, pp. 35–42, 2020.
- [13] Arikunto, Dasar - Dasar Evaluasi Pendidikan. Jakarta: Bumi Aksara, 2007.
- [14] D. Kurniadi, Y.H. Augustin, H.I N. Akbar and I. Farida, “Penerapan Algoritma k-Means Clustering untuk Pengelompokan Pembangunan Jalan pada Dinas Pekerjaan Umum dan Penataan Ruang”, Jurnal Teknologi Informasi (AITI), Vol. 20, No. 1, pp 64-77, Februari 2023.
- [15] J. Suntoro, “Data Mining Algoritma dan Implementasi dengan Pemrograman PHP,” Elex Media Komputindo., 2019
- [16] A. T. Rahman, “Coal Trade Data Clustering Using K-Means

- (Case Study Pt. Global Bangkit Utama)”, ITSMART J. Teknol. dan Inf., vol. 6, no. 1, pp. 24–31, 2017.
- [17] D.A.I.C. Dewi and D.A.K. Pramita, “Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma *Clustering* K-Medoids dalam Pengelompokan Produksi Kerajinan Bali”, *Jurnal Matrix*, Vol. 9, No. 3, November 2019.
- [18] R.A. Candra, Y.H. Chrisnanto and P.N. Sabrina, “Segmentasi Mahasiswa Berdasarkan Karakteristik Pola Belajar Menggunakan Metode K-Medoids Clustering”, *Prosiding Seminar Nasional Sains dan Teknologi*, Vol. 12, No. 1, pp 355-361, 2022.
- [19] M. Iqbal, Syaripuddin and M.N. Huda, “, Implementasi Algoritma K-Means Clustering dengan Jarak Euclidean dalam Mengelompokkan Daerah Penyebaran COVID-19 di Kabupaten Bogor”, *Jurnal Ilmiah Mat*