

## PENERAPAN ALGORITMA *K-MEANS* UNTUK ANALISIS DATA ULASAN DI SITUS TRIPADVISOR

Maya Nurjanah<sup>1</sup>, Toni Arifin<sup>2</sup>

<sup>1</sup>Universitas Adhirajasa Reswara Sanjaya  
e-mail: mayanurj89@gmail.com

<sup>2</sup>Universitas Adhirajasa Reswara Sanjaya  
e-mail: toni.arifin@ars.ac.id

### Abstrak

Ulasan atau *review* merupakan suatu pendapat langsung dari seseorang. *Review* yaitu beberapa dari salah satu faktor yang menentukan keputusan seseorang, dengan menunjukkan bahwa orang dapat mengambil jumlah *review* sebagai indikator popularitas tempat atau nilai dari suatu tempat dengan mempengaruhi kemauan untuk mendatangi suatu tempat pariwisata. Penelitian ini membahas tentang Penerapan metode *K-Means* untuk *review* data travel . Dataset ini diambil langsung melalui halaman situs web *UCI Machine Learning Repository* dengan nama *Review Travel* dengan jumlah data sebanyak 980 record, terdiri dari 10 variabel atau atribut yaitu galeri seni, klub dansa, bilah jus, restoran, museum, *resort*, taman atau tempat piknik, pantai, bioskop dan lembaga keagamaan. Proses *cluster* dibagi kedalam 2 (dua) *cluster* yaitu *cluster rating* tertinggi (C0) dan *cluster rating* terendah (C1). Sehingga diperoleh hasil pengelompokkan data ulasan atau *review rating* wisatawan *cluster* tertinggi berada pada C0 terdiri dari 3 objek wisata unggulan dengan *rating* tertinggi pada C0 yaitu Galeri seni dengan *rating* = 7,118, Museum = 5,396 dan Klub dansa = 4,783. Sedangkan *cluster* terendah berada pada C1 terdiri dari 3 objek wisata unggulan dengan *rating* tertinggi yaitu Restoran dengan *rating* = 10,394, Klub dansa = 3,361, Galeri seni = 3,233.

**Kata Kunci:** Data Mining, *Review Travel*, *Clustering*, *K-Means*.

### Abstract

A review or review is a direct opinion from someone. Reviews are some of the factors that determine a person's decision, by showing that people can take the number of reviews as an indicator of the popularity of a place or the value of a place by influencing the willingness to visit a tourism place. This study discusses the application of the *K-Means* method for reviewing travel data. This dataset is taken directly through the *UCI Machine Learning Repository* website page with the name *Travel Review* with a total data amount of 980 records, consisting of 10 variables or attributes, namely art galleries, dance clubs, juice bars, restaurants, museums, resorts, parks or picnic spots, beaches, cinemas and religious institutions. The cluster process is divided into 2 (two) clusters, namely the highest cluster rating (C0) and the lowest cluster rating (C1). So that the results of grouping data reviews or reviews the highest tourist rating cluster is at C0 consisting of 3 leading tourist objects with the highest rating at C0, namely art galleries with *rating* = 7,118, museum = 5,396 and dance club = 4,783. While the lowest cluster is at C1 consisting of 3 leading tourist objects with the highest rating, namely restaurants with *rating* = 10.394, dance clubs = 3.361, art galleries = 3.233.

**Keywords:** Data Mining, *Travel Review*, *Clustering*, *K-Means*.

## 1. Pendahuluan

Pariwisata merupakan salah satu faktor yang sangat berpengaruh pada perkembangan perekonomian suatu daerah yang didalamnya memiliki tujuan dengan meningkatkan devisa negara, meningkatkan pendapatan pada suatu daerah, memperluas pada suatu kesempatan kerja, dan tidak mengeksploitasi kekayaan suatu alam (Dewi, Herdiani, & Kusumo, 2018).

Wisatawan harus mengetahui terlebih dahulu kondisi pada objek wisata tersebut, sebelum para wisatawan ingin berkunjung ke suatu tempat objek wisata. Hal ini dapat dilihat dari ulasan atau *review* wisatawan yang sebelumnya pernah mengunjungi suatu tempat (R. Sari & Hayuningtyas, 2019). Sehingga dapat diartikannya wisatawan adalah orang – orang yang berkunjung pada suatu tempat atau dengan maksud tujuan tertentu (R. W. Sari & Hartama, 2018).

Perkembangan yang sangat pesat pada media *online*, tidak jarang wisatawan membagikan pengalaman dalam berwisata mereka pada situs web yang didalamnya menyediakan kolom komentar pada suatu tempat wisata, seperti pada situs web Tripadvisor. Situs web wisata seperti TripAdvisor cukup berperan penting bagi wisatawan sebagai salah satu bahan pertimbangan untuk mengunjungi suatu tempat yang diinginkan (Dewi et al., 2018).

Dengan beberapa situs yang muncul untuk memberikannya kemudahan bagi konsumen untuk memberikan ulasan (*review*) atau penilaian terhadap suatu entitas (Maulana & Sibaroni, 2019).

Jika dilihat dari rata-rata ulasan atau *review* wisatawan pada tahun 2018. Khususnya ulasan rating wisatawan ke suatu objek wisata di negara asia timur. Tercatat ada 10 (sepuluh) objek wisata seperti Galeri seni, Klub dansa, Bilah jus, Restoran, Museum, Resort, Taman / Tempat piknik, Pantai, Bioskop dan Lembaga Keagamaan (Maulida, 2018).

Gambaran dari kepuasan konsumen untuk memilih suatu tempat. Banyaknya data yang tak terhingga dari berbagai konsumen dan beberapa ulasan atau *review* juga belum ditetapkan adanya *rating* maka membuat pihak travel mengalami kesulitan dalam mengklasifikasikan ulasan atau *review* kedalam *rating* dan mengetahui kekurangan dari tempat tersebut. Sehingga sangat penting adanya sistem yang membantu untuk mempercepat prediksi

*rating* berdasarkan ulasan atau *review* wisatawan supaya tempat tersebut dapat melakukan evaluasi untuk ke depannya menjadi lebih baik. (Paramitha & Dewi, 2019).

Penelitian terdahulu yang meneliti tentang pengelompokan objek wisata unggul yang berbeda pada suatu tempat. Dimana perbedaannya terletak pada suatu objek penelitian yang akan diteliti. Tujuannya penelitian ini adalah untuk mengetahui suatu objek wisata unggulan di negara asia timur yang memiliki potensi yang paling tinggi dengan wisatawan memberikan ulasan berupa *rating* ke masing – masing tempat sehingga menemukan tempat yang unggul dengan peringkat yang tinggi. Pengelompokan tersebut dapat menggunakan metode pengelompokan dengan algoritma *K-Means*. Dengan data yang sudah dikelompokkan menggunakan algoritma *K-Means* diharapkan dapat mempermudah dalam mengelompokkan *rating* wisatawan dengan memberikan ulasan ke suatu objek wisata unggulan dengan peringkat yang tinggi, sehingga dapat melakukan perbaikan sarana dan prasarana objek wisata unggulan yang dapat meningkatkan *rating* wisatawan pada objek wisata (Maulida, 2018).

Untuk memberikan rekomendasi suatu tempat berdasarkan ulasan atau *review* wisatawan, maka digunakanlah metode *K-Means*. Dengan metode tersebut, diharapkan akan menghasilkan pengelompokan objek wisata unggul pada suatu tempat (Pratama, Sari, & Adikara, 2018).

*Clustering* adalah salah satu teknik dari salah satu fungsionalitas data mining, algoritma *clustering* adalah pengelompokan dari sejumlah data menjadi suatu kelompok – kelompok data tertentu (Silalahi, 2018).

Untuk meringkas data atau sejumlah variabel menjadi lebih sedikit analisis *cluster* ini menjadi sangat berguna. Pada proses data ini dapat dilakukan dengan mengelompokkan objek – objek berdasarkan kesamaan karakteristik tertentu diantara objek – objek yang akan diteliti (Anshori & Nuraini, 2020).

Tujuan analisis *cluster* ini untuk mengelompokkan / mengklasifikasikan objek – objek berdasarkan karakteristik berdasarkan objek – objek, sehingga diketahui ciri khas dari masing – masing kelompok (Ary, 2015).

Berdasarkan pemaparan latar belakang permasalahan, maka pada kesempatan kali ini akan dilakukan penelitian terhadap analisis data ulasan pada situs Tripadvisor. Data tersebut akan dikelompokkan dengan metode *cluster* dan algoritma *K-Means* dengan hasil yang diharapkan adalah berupa informasi mengenai pengelompokan data ulasan berdasarkan rating tertinggi wisatawan pada suatu objek tempat wisata.

## 2. Metode Penelitian

Bagian ini menjelaskan tahapan demi tahapan yang penulis lakukan, dalam menyusun laporan ini agar lebih sistematis dan terarah sesuai dengan yang telah direncanakan.

### Tahapan Penelitian

#### 1. Pengumpulan data

Dalam tahap ini akan dijelaskan bagaimana data tersebut dikumpulkan dan darimana data tersebut diperoleh sehingga data tersebut dapat digunakan dalam eksperimen penelitian.

#### 2. Pengolahan awal data

Pada tahap ini akan dijelaskan bagaimana pengolahan data untuk mendapatkan data yang sudah ditransformasikan agar sesuai dengan bentuk yang diinginkan.

#### 3. Model yang diusulkan

Setelah dilakukan tahap pengolahan data pada tahap sebelumnya.

#### 4. Eksperimen dan pengujian model

Selanjutnya dalam tahapan ini akan dilakukan pengujian terhadap model yang diusulkan untuk mendapatkan hasil kinerja dari model yang diusulkan.

#### 5. Evaluasi dan validasi hasil

Pada tahapan penelitian terakhir ini akan dilakukan evaluasi terhadap eksperimen dan pengujian pada model yang diusulkan sehingga dapat mengetahui hasil kinerja dalam penelitian ini.

### Pengumpulan Data

Dataset ini diambil langsung melalui halaman situs *web UCI Machine Learning Repository* dengan nama *Review Travel* dengan jumlah data sebanyak 980 record, terdiri dari 10 variabel atau atribut, dan variabel tujuan yaitu variabel yang dijadikan sebagai ulasan atau *review* wisatawan dengan memberikan nilai berupa *rating* pada suatu tempat. Variabel atau atributnya yaitu galeri seni, klub dansa, bilah jus, restoran, museum, *resort*, taman

atau tempat piknik, pantai, bioskop dan lembaga keagamaan.

Tabel 1.

Deskripsi atribut dataset *Review Travel*

No	Atribut	Deskripsi	Tipe Data
1	Galeri Seni	Umpan balik pengguna rata-rata galeri seni	Numerikal
2	Klub Dansa	Umpan balik pengguna rata-rata klub dansa	Numerikal
3	Bilah Jus	Umpan balik pengguna rata-rata bilah jus	Numerikal
4	Restoran	Umpan balik pengguna rata-rata restoran	Numerikal
5	Museum	Umpan balik pengguna rata-rata museum	Numerikal
6	Resort	Umpan balik pengguna rata-rata resort	Numerikal
7	Taman/Tempat piknik	Umpan balik pengguna rata-rata taman/tempat piknik	Numerikal
8	Pantai	Umpan balik pengguna rata-rata pantai	Numerikal
9	Bioskop	Umpan balik pengguna rata-rata bioskop	Numerikal
10	Lembaga Keagamaan	Umpan balik pengguna rata-rata lembaga keagamaan	Numerikal

### Pengolahan Awal Data

Pada penelitian ini menggunakan dataset yang di dapatkan dari halaman situs *web UCI Machine Learning Repository* yaitu berupa dataset *review travel* sebanyak 980 *record*. Data yang sudah dijadikan sampel akan dilakukan pengolahan data dengan proses *clustering* dengan menggunakan algoritma *K-Means* sehingga didapatkanlah hasil pengelompokan data yang diinginkan. Adapun langkah dalam *cluster* dengan algoritma *K-Means* yaitu :

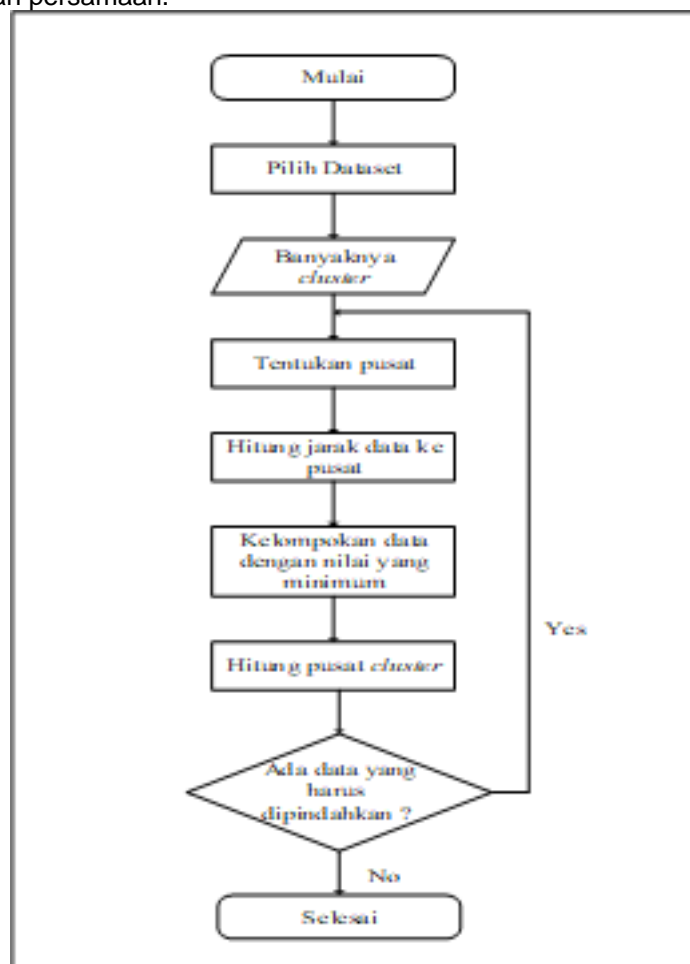
1. Menentukan pusat *cluster* awal. Menentukan *centroid* awal dilakukan secara acak dari data/objek yang tersedia sebanyak jumlah *cluster* k. Nilai *centroid* awal pada penelitian ini dilakukan pemilihan secara acak.
2. Menghitung jarak dengan pusat cluster Untuk menghitung jarak setiap data yang ada terhadap pusat cluster dalam

- penelitian ini penulis menggunakan rumus *Euclidean Distance*.
3. Pengelompokan data Alokasikan masing-masing data ke dalam *centroid* yang paling terdekat. Dalam mengalokasikan kembali data ke dalam masing-masing *cluster* didasarkan pada perbandingan jarak antara data dengan *centroid* setiap *cluster* yang ada, data dialokasikan secara tegas ke dalam *cluster* yang mempunyai jarak ke *centroid* terdekat dengan data tersebut. Berikut ini merupakan hasil perbandingan jarak antara data dengan *centroid* setiap *cluster* yang ada berdasarkan perhitungan jarak dengan *Euclidean Distance* untuk iterasi 0.
  4. Lakukan iterasi, kemudian tentukan posisi *centroid* baru dengan menggunakan persamaan.
  5. Ulangi langkah 3 jika posisi *centroid* baru tidak sama. Pengecekan konvergensi dilakukan dengan membandingkan *matriks group assignment* pada iterasi sebelumnya dengan *matrik group assignment*.
  6. pada iterasi yang sedang berjalan. Jika hasilnya sama maka algoritma *kmeans cluster analysis* sudah *konvergen*, tetapi jika berbeda maka belum *konvergen* sehingga perlu dilakukan iterasi berikutnya.

### 3. Hasil dan Pembahasan

Pada bagian ini akan dijelaskan pembahasan mengenai hasil penelitian yang penulis lakukan.

#### 3.1. Perhitungan Algoritma K-Means



Gambar 1. Model yang Diusulkan

Langsung melalui halaman situs web *UCI Machine Learning Repository* dengan nama *Review Travel*. Sample data sebanyak 100 record dari laporan *review* atau ulasan *rating*

wisatawan dengan memberikan *rating* ke suatu objek wisata unggulan.

Tahap pertama dalam algoritma *K-Means* adalah jumlah *cluster*, pada

penelitian ini terdapat 2 *cluster*. Penentuan *cluster* awal dilakukan secara *random* dengan 10 atribut yaitu galeri seni, klub dansa, billah jus, restoran, museum, resort,

taman / tempat piknik, pantai dan lembaga keagamaan. Berikut merupakan datasetnya terdapat pada Tabel 2.

Tabel 2.  
Jumlah Ulasan / *review rating* wisatawan

No	Galeri Seni	Klub Dansa	Billah Jus	Restoran	Museum	Resort	Taman/Tempat Piknik	Pantai	Bioskop	Lembaga keagamaan
1	0,93	1,8	2,29	0,62	0,8	2,42	3,19	2,79	1,82	2,42
2	1,02	2,2	2,66	0,64	1,42	3,18	3,21	2,63	1,86	2,32
3	1,22	0,8	0,54	0,53	0,24	1,54	3,18	2,8	1,31	2,5
4	0,45	1,8	0,29	0,57	0,46	1,52	3,18	2,96	1,57	2,86
5	0,51	1,2	1,18	0,57	1,54	2,02	3,18	2,78	1,18	2,54
6	0,99	1,28	0,72	0,27	0,74	1,26	3,17	2,89	1,66	3,66
7	0,9	1,36	0,26	0,32	0,86	1,58	3,17	2,66	1,22	3,22
8	0,74	1,4	0,22	0,41	0,82	1,5	3,17	2,81	1,54	2,88
9	1,12	1,76	1,04	0,64	0,82	2,14	3,18	2,79	1,41	2,54
10	0,7	1,36	0,22	0,26	1,5	1,54	3,17	2,82	2,24	3,12
...	...	...	...	...	...	...	...	...	...	...
90	0,67	2,08	1,04	0,54	0,74	1,64	3,18	2,86	1,57	2,64
91	1,38	1,12	0,72	0,53	0,88	1,36	3,18	2,81	1,79	3,26
92	0,56	1,16	0,13	0,49	0,54	0,9	3,17	2,91	1,54	2,64
93	1,1	2,04	2,45	0,54	1,44	2,28	3,19	2,74	1,63	2,42
94	0,64	1,2	2,27	0,64	1,42	2,83	3,2	2,79	1,47	2,38
95	0,99	2	2,13	0,64	1,82	2,6	3,19	2,67	1,5	2,46
96	1,12	0,96	0,3	0,46	0,56	1,92	3,18	2,98	1,22	2,96
97	0,9	1,44	2,34	0,55	0,94	2,14	3,19	2,79	1,79	2,82
98	0,9	0,56	1,38	0,37	0,8	1,9	3,18	2,91	1,41	2,54
99	1,02	0,886	2,7	0,53	1,02	1,78	3,19	2,86	2,14	2,78
100	0,67	1,6	0,21	0,47	2,16	1,96	3,17	2,86	3,04	3,1

Keterangan :

Rating 4 = Sangat Baik  
Rating 3 = Sangat Bagus  
Rating 2 = Rata – Rata  
Rating 1 = Buruk  
Rating 0 = Mengerikan

Tahap selanjutnya yaitu menentukan titik pusat awal dari *cluster* (*centroid*) yang dipilih secara *random*. Pada penelitian ini diambil dari data ke 20 dan 57.

Tabel 3. *Centroid* Awal

Centroid	Galeri seni	Klub Dansa	Billah Jus	Restoran	Museum	Resort	Taman/Tempat piknik	Pantai	Bioskop	Lembaga Keagamaan
C0	0,8	1,04	2,1	0,58	1,18	1,98	3,19	2,93	1,22	2,48
C1	0,8	1,16	2,43	0,54	1,06	2,28	3,19	2,86	1,44	2,48

Setelah menentukan *centroid* awal, maka tahap selanjutnya yaitu menghitung jarak setiap data yang ada ke *centroid* terdekat untuk menentukan *cluster* menggunakan rumus *Euclidean distance*. Berikut contoh perhitungan jarak ke setiap *centroid*

Data ke-1 cluster ke 0

$$d(xic0) = \sqrt{a^2 + b^2}$$

$$= \sqrt{(0,93 - 0,8)^2 + (1,8 - 1,04)^2 + (2,29 - 2,1)^2 + (0,62 - 0,58)^2 + (0,8 - 1,18)^2 + (2,42 - 1,98)^2 + (3,19 -$$

$$3,19)^2 + (2,79 - 2,93)^2 + (1,82 - 1,22)^2 + (2,42 - 2,48)^2}$$

$$= \sqrt{0,13^2 + 0,76^2 + 0,19^2 + 0,04^2 + 0,38^2 + 0,44^2 + 0 + 0,14^2 + 0,6^2 + 0,06^2}$$

$$= \sqrt{0,0169 + 0,5776 + 0,0361 + 0,0016 + 0,1444 + 0,1936 + 0 + 0,0196 + 0,36 + 0,0036}$$

$$= \sqrt{1,3534} = 1,163$$

Dari perhitungan data ke-1 dengan *cluster* ke-0 maka dihasilkan jarak 1,163.

**Data ke-1 cluster ke 1**

$$\begin{aligned}
&= \sqrt{(0,93 - 0,8)^2 + (1,8 - 1,16)^2 + (2,29 - 2,43)^2 + (0,62 - 0,54)^2 + (0,8 - 1,06)^2 + (2,42 - 2,28)^2} \\
&+ \sqrt{(3,19 - 3,19)^2 + (2,79 - 2,86)^2 + (1,82 - 1,44)^2 + (2,42 - 2,48)^2} \\
&= \sqrt{0,13^2 + 0,64^2 + (-0,14^2) + 0,08^2 + (-0,26^2) + 0,14^2 + 0 + (-0,07^2) + 0,38^2 + (-0,06^2)} \\
&= \sqrt{0,0169 + 0,4096 + 0,0196 + 0,0064 + 0,0676 + 0,0196 + 0 + 0,0049 + 0,1444 + 0,0036} \\
&= \sqrt{0,6926} = 0,832
\end{aligned}$$

Dari perhitungan data ke-1 dengan *cluster* ke-1 maka dihasilkan jarak 0,832.

Hasil perhitungan tersebut yaitu jarak data perhitungan tersebut yaitu jarak data ke-1 dengan *cluster* ke-0 sebesar 1,163 jarak dan data ke-1 dengan *cluster* ke-1 sebesar 0,832. Berikut Tabel IV.4 yang merupakan perhitungan lengkap dari iterasi ke-1.

Tabel 4.  
Hasil perhitungan dengan menggunakan rumus *Euclidean* pada iterasi ke-1

Data ke-i	Jarak ke centroid		Jarak Terdekat	Cluster yang diikuti
	C0	C1		
1	1,1633572 11	0,8322259 31	0,83222 5931	C1
2	1,9324595 73	1,5452507 9	1,54525 079	C1
3	1,9423696 87	2,2625649 16	1,94236 969	C0
4	2,2302914 61	2,4943135 33	2,23029 146	C0
5	1,0560303 03	1,4229546 73	1,05603 03	C0
6	2,0962347 2	2,3732677 89	2,09623 472	C0
7	2,1092415 7	2,4441153 82	2,10924 157	C0
8	2,0817540 68	2,4085472 8	2,08175 407	C0
9	1,4009282 64	1,5783535 73	1,40092 826	C0
10	2,3468489 51	2,6086969 93	2,34684 895	C0
...	...	...	...	..
90	1,6389020 72	1,8303005 22	1,63890 207	C0
91	1,9162202 38	2,2078269 86	1,91622 024	C0
92	2,3802940 99	2,7501636 31	2,38029 41	C0
93	1,2567816 04	1,0311643 9	1,03116 439	C1
94	0,9777525 25	0,7147027 35	0,71470 274	C1

95	1,3793476 72	1,2497199 69	1,24971 997	C1
96	1,9913312 13	2,3108007 27	1,99133 121	C0
97	0,8785214 85	0,6112282 72	0,61122 827	C1
98	0,9969453 35	1,3116783 14	0,99694 533	C0
99	1,1867181 64	1,0080178 57	1,00801 786	C1
100	2,9287540 01	3,0661050 21	2,92875 4	C0

Selanjutnya kelompokan data sesuai dengan jarak *cluster* terdekat. Dari data yang sudah dikelompokan akan didapat *centroid* baru dari hasil rata – rata setiap *cluster*. Lakukan langkah tersebut sampai nilai *centroid* tidak mengalami perubahan.

**3.2. Validasi Hasil**

Berdasarkan Uji DBI dapat ditentukan seberapa baik *clustering* sudah dilakukan dengan parameter semakin kecil atau minimal nilai uji maka hasil *clustering* yang didapat semakin baik. Setelah dilakukan percobaan menggunakan *Rapidminer* dengan pemodelan K=2 sampai dengan K=10, didapatkan nilai *cluster* optimal terdapat pada K=2 dengan nilai DBI 0,101.

Tabel 5.  
Hasil DBI pengujian K2 sampai K10 pada aplikasi *Rapidminer*

Cluster	Hasil
K=2	0,101
K=3	0,147
K=4	0,164
K=5	0,171
K=6	0,175
K=7	0,171
K=8	0,172
K=9	0,165
K=10	0,172

Pada penentuan *cluster* menggunakan matriks *Devies Bouldin Index* digunakan data centroid dari iterasi terakhir pada proses *clustering*. Berikut merupakan centroid akhir dari hasil *clustering*.

**3.3. Pembahasan**

*Rapidminer*, maka menghasilkan informasi sebagai berikut :

1. Pada *cluster* 0 terdiri dari 959 items dengan tempat pariwisata yang paling unggul di *cluster* 0 dengan *rating* tertinggi di masing – masing tempat, diantaranya Galeri seni dengan *rating* =

- 7,118, Klub dansa = 4,783, Billah jus = 3,305, Restoran = 2,994, Museum = 5,396, *Resort* = 3,553, Tempat / tempat piknik = 3,714, Pantai = 4,036, Bioskop = 4,390, Lembaga keagamaan = 2,678.
2. Pada *cluster* 1 terdiri dari 21 items dengan tempat pariwisata yang paling unggul di *cluster* 1 dengan *rating* tertinggi di masing – masing tempat, diantaranya Galeri seni dengan *rating* = 3,233, Klub dansa = 3,361, Billah jus = 1,353, Restoran = 10,394, Museum = 2,470, *Resort* = 2,033, Tempat / tempat piknik = 2,436, Pantai = -1,564, Bioskop = 1,318, Lembaga keagamaan = 0,749. Berdasarkan uraian tersebut, maka dapat disimpulkan bahwa *cluster* tertinggi berada pada *cluster* 0. Sehingga tempat yang menjadi unggul dengan *rating* tertinggi pada *cluster* 0 yaitu Galeri seni, Museum, dan Klub dansa. Sedangkan *cluster* terendah berada pada *cluster* 1. Sehingga tempat yang menjadi unggul dengan *rating* tertinggi pada *cluster* 1 yaitu Restoran, Klub dansa, dan Galeri seni.

#### 4. Kesimpulan

1. Untuk mengetahui suatu data ulasan atau *review* dengan *rating* tertinggi pada suatu tempat pariwisata maka dapat menerapkan metode *clustering K-Means*, data diolah menggunakan *Ms. Excel* untuk ditentukan nilai *centroid* dalam 2 *cluster* yaitu *cluster rating* tertinggi (C0) dan *cluster rating* terendah (C1).
2. Berikut hasil pengelompokan data ulasan atau *review rating* wisatawan C0 terdiri dari 3 objek wisata unggulan dengan *rating* tertinggi pada *cluster* 0 yaitu Galeri seni dengan *rating* = 7,118, Museum = 5,396 dan Klub dansa = 4,783. Dan pada *cluster* 1 terdiri dari 3 objek wisata unggulan dengan *rating* tertinggi pada *cluster* 1 yaitu Restoran dengan *rating* = 10,394, Klub dansa = 3,361, Galeri seni = 3,233.

#### Saran

1. Eksperimen *review* travel dapat menggunakan algoritma *clustering* yang

lain seperti *K-Medoids* dan *Fuzzy C-Means* sehingga mendapatkan hasil yang lebih baik.

2. Dapat mengembangkan penelitian dengan menambahkan atribut waktu seperti tanggal dan bulan untuk penambahan atribut yang dijadikan untuk mengetahui waktu ulasan atau *review rating* wisatawan pada dataset *review* travel.

#### Referensi

- Anshori, I. F., & Nuraini, Y. (2020). Pengelompokan Data Kecelakaan Lalu Lintas Di Kota Tasikmalaya Menggunakan Algoritma K-Means. *2*(1), 118–127.
- Ary, M. (2015). Pengklasifikasian Karakteristik Mahasiswa Baru Dalam Memilih Program Studi Menggunakan Analisis Cluster. *II*(1), 181–188.
- Dewi, Herdiani, & Kusumo, D. S. (2018). Multi-Aspect Sentiment Analysis Komentar Wisata TripAdvisor dengan Rule-Based Classifier ( Studi Kasus : Bandung Raya ). *E-Proceeding of Engineering*, *5*(1), 1589–1596.
- Maulana, A., & Sibaroni, Y. (2019). Klasifikasi Sentimen Ulasan Tempat Makan Berbahasa Indonesia Menggunakan Lexicon dan Improved Naive Bayes. *Ind. Journal on Computing*, *4*(3), 107–116. <https://doi.org/10.21108/indojc.2019.4.3.369>
- Maulida, L. (2018). Kunjungan Wisatawan Ke Objek Wisata Unggulan Di Prop . DKI Jakarta Dengan K-Means. *JISKa (Jurnal Informatika Sunan Kalijaga)*, *2*(3), 167–174.
- Pratama, N. D., Sari, Y. A., & Adikara, P. P. (2018). Analisis Sentimen Pada Review Konsumen Menggunakan Metode Naive Bayes Dengan Seleksi Fitur Chi Square Untuk Rekomendasi Lokasi Makanan Tradisional. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *2*(9), 2982–2988.
- Paramitha, D. T. A., & Dewi, I. C. & C. (2019). Prediksi Rating Otomatis Berdasarkan Review Restoran pada Aplikasi Zomato dengan menggunakan Extreme Learning

- 
- Machine ( ELM ). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(5), 4687–4693.
- Sari, R., & Hayuningtyas, R. Y. (2019). Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Pada Wisata TMII Berbasis Website. *IJSE – Indonesian Journal on Software Engineering*, 5(2), 51–60.
- Sari, R. W., & Hartama, D. (2018). Data Mining : Algoritma K-Means Pada Pengelompokan Wisata Asing ke Indonesia Menurut Provinsi. 322–326.
- Silalahi, M. (2018). Analisis Clustering Menggunakan Algoritma K-Means Terhadap Penjualan Produk Pada PT. Batamas Niaga Jaya. *Computer Based Information System Journal*, 06(2), 20–35.