
Penerapan Algoritma K-Means Terhadap Kunjungan Wisatawan Asing Di Hotel Berbintang di Indonesia

Corie Mei Hellyana^{1*}, Nuzul Imam Fadlilah²

¹ *Sistem Informasi Akuntansi, Universitas Bina Sarana Informatika, Indonesia*

² *Teknologi Komputer, Universitas Bina Sarana Informatika, Indonesia*

E-mail:^{1}corie.cma@bsi.ac.id, ²nuzul.nfh@bsi.ac.id*

Abstrak

Sebagai negara kepulauan, Indonesia memiliki banyak daya tarik wisata sehingga banyak tamu asing atau wisatawan yang berkunjung ke daerah di Indonesia. Dengan adanya daya tarik wisata, pemerintah berupaya untuk menyediakan fasilitas hotel guna menunjang dan meningkatkan ekonomi masyarakat. Hotel sendiri merupakan tempat khusus yang disediakan untuk menginap yang dikelola oleh pihak swasta. Tujuan dari penelitian ini adalah untuk mengetahui pengelompokan tamu asing yang menginap di hotel berbintang berdasarkan propinsi mulai tahun 2016 sampai tahun 2021. Perhitungan pengelompokan menggunakan algoritma K-Means, dimana algoritma K-Means adalah satu jenis metode pengelompokan non-hierarki dengan membagi data dengan jumlah yang besar menjadi cluster/kelompok. Hasil yang didapat dari implementasi algoritma K-Means ini didapatkan 4 propinsi masuk ke dalam kelompok/cluster 0 (jumlah kunjungan tinggi), 6 propinsi masuk ke dalam cluster 1 (jumlah kunjungan sedang) dan 24 propinsi masuk ke dalam cluster 2 (jumlah kunjungan rendah).

Kata kunci: clustering, k-means, tamu asing, hotel berbintang

Abstract

As an island country, Indonesia has many tourist attractions, so many foreign guests or tourists visit areas in Indonesia. With tourism attraction, the government wants to provide hotel facilities to support and improve the community's economy. The hotel itself is a special privately run property. The aim of this study is to find out the grouping of foreign guests in star hotels from 2016 to 2021 by province. The clustering calculation uses the K-Means algorithm, where the K-Means algorithm is a type of non-hierarchical clustering method that divides large data into clusters/groups. The results obtained from the application of the K-Means algorithm show that there are 4 provinces in group/cluster 0 (high number of visits), 6 provinces in cluster 1 (moderate number of visits) and 24 provinces in cluster 2 (low number of visits).

Keywords: clustering, k-means, foreign guest, star hotel

1. Pengantar

Sebagai salah satu negara kepulauan terbesar di dunia, Indonesia mempunyai banyak tempat wisata seperti wisata bahari, wisata gunung dan taman atau situs cagar budaya. Dengan adanya destinasi wisata ini, sedikit banyak akan menarik minat wisatawan, baik domestik maupun mancanegara untuk berkunjung ke Indonesia sehingga meningkatkan pendapatan masyarakat di lingkungan destinasi wisata tersebut. Apabila jumlah wisatawan yang berkunjung semakin bertambah, maka semakin banyak pula potensi bisnis yang ditawarkan oleh masyarakat, antara lain tempat makan, penginapan, transportasi dan lainnya.

Menurut Peraturan Pemerintah No. 56 Tahun 2011 menyebutkan bahwa Hotel merupakan bangunan yang secara khusus digunakan khusus untuk menginap atau beristirahat dengan mendapatkan fasilitas dan pelayanan oleh pihak pengelola secara berbayar (Sari & Hartama, 2018).

Sedangkan pengertian Hotel berdasarkan (Direktur Jenderal Pariwisata Nomor: 14/U/II/88, 1988) yaitu jenis penginapan yang menggunakan sebagian/bangunan atau sebagiannya untuk menyediakan tempat tinggal, makan dan minum. layanan dan utilitas umum lainnya yang dikelola secara komersial. Hotel memiliki beberapa tingkatan/taraf yakni

hotel berbintang 1, hotel berbintang 2, hotel berbintang 3, hotel berbintang 4 dan hotel berbintang 5.

Dalam penelitian ini, informasi data jumlah kunjungan wisatawan asing pada hotel bintang menurut propinsi yang didapatkan dari laman Badan Pusat Statistik (BPS). BPS sendiri merupakan lembaga pemerintah non-departemen yang berada dibawah langsung dan bertanggung jawab kepada Presiden (Indonesia, 2007). Adanya penelitian ini diharapkan ada upaya untuk lebih mengembangkan daya tarik wisata pada daerah/propinsi yang masih memiliki jumlah tamu asing sedikit dibandingkan dengan daerah/propinsi lain.

Penelitian dengan judul Analisis Algoritma K-Medoids Clustering Dalam Menentukan Pemesanan Hotel (Kurniawan et al., 2020) menghasilkan tingkat pemesanan hotel tertinggi ada pada cluster 2 dengan 4530 item/pemesanan terbanyak. Dalam penelitian (Sari & Hartama, 2018) wisata asing dikelompokkan menjadi 3 cluster menggunakan algoritma K-Means dengan rincian cluster 0 sejumlah 1 propinsi, cluster 1 sejumlah 2 propinsi dan cluster 2 sejumlah 31 propinsi. (Listiani et al., 2019) melakukan penelitian dalam mengelompokkan data penduduk bungursari ke dalam 4 cluster yakni C1 untuk pekerja mandiri, C2 untuk pekerjaan wirausaha, C3

untuk pekerja swasta dan C4 untuk pekerja paruh waktu dengan nilai akurasi 79%.

UU Kepariwisata No. 10 Tahun 2009 mendefinisikan pengunjung atau turis asing sebagai orang yang melakukan perjalanan dekat atau jauh dengan tujuan mencari kesenangan, mencari informasi atau bepergian ke luar tempat tinggalnya (Pramayuda et al., 2020). Dalam prosesnya wisatawan atau tamu asing ini terbagi menjadi dua yakni wisatawan dalam negeri dan wisatawan luar negeri atau wisatawan asing.

Berdasarkan paparan penelitian sebelumnya, penulis mencoba melakukan penelitian guna mempersiapkan strategi kedepan bagi pihak hotel berbintang dalam menyambut tamu asing yang akan menginap.

Teknologi data mining sangat bermanfaat dalam menemukan informasi penting. Data mining dapat digunakan dalam proses mengidentifikasi, mengelompokkan, memprediksi, mengestimasi atau bahkan menentukan aturan-aturan asosiasi yang sudah ada (Kurniawan et al., 2020). Dengan menggunakan metode clustering yang bekerja dengan cara membagi data berdasarkan ciri dari kelompok masing-masing kemudian mendistribusikan dalam kelompok sehingga didapatkan data

beberapa tingkat yang saling berhubungan (Nasir, 2021).

K-Means adalah algoritma pengelompokan iteratif. Ketika sebuah nilai ditugaskan ke sebuah cluster, itu dilakukan secara acak atau random, dengan nilai yang menjadi nilai tengah disebut centroid, atau mean. Algoritma yang paling mudah diterapkan adalah algoritma K-Means karena dapat mengolah data yang besar secara tepat waktu, efisien dan efektif. Kelemahan K-Means terletak pada penentuan cluster awal yang bergantung pada input data yang diberikan.

Pengelompokan data dengan menggunakan algoritma K-Means melalui beberapa tahapan perhitungan, yaitu (Purba et al., 2019):

- a. Menentukan nilai k , untuk banyaknya jumlah cluster yang diinginkan,
- b. Memilih sebanyak k data dari set data secara random untuk dijadikan sebagai pusat cluster (centroid) awal,
- c. menentukan jarak objek dengan pusat data masing-masing, yang menggunakan rumus *Euclidean Distance*,
- d. Mengelompokkan objek berdasarkan jarak terdekat dengan pusat data,
- e. Menentukan pusat data (centroid) terbaru dengan cara menghitung rata-rata nilai data tiap kelompok,

- f. Lakukan perulangan untuk langkah 3 dan 4, sampai dengan tidak ada lagi objek yang berpindah cluster.

Rumus yang digunakan untuk menentukan nilai centroid yaitu (Khotimah, 2014) dalam:

$$v_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} x_{ki} \quad (1)$$

Dimana:

v_{ij} = pusat data rata-rata cluster ke-i untuk variabel ke-j

N_i = jumlah anggota cluster ke-i

i, k = indeks dari cluster

j = indeks dari variabel

x_{kj} = nilai data ke-k variabel ke-j dalam cluster

Menurut (Atma et al., 2022) dalam menghitung nilai korelasi antar dua objek, digunakan rumus *Euclidean Distance* sebagai berikut:

$$d_{euclidean}(x, y) = \sqrt{(x_i - y_i)^2} \quad (2)$$

Dimana:

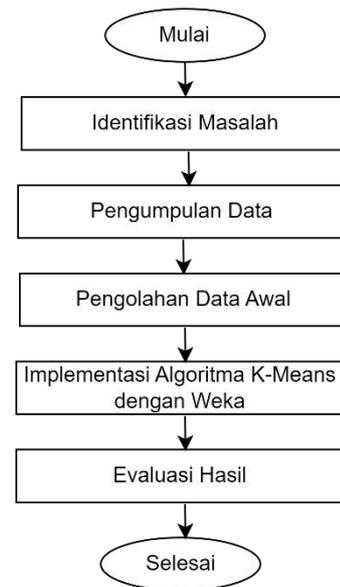
$d(x,y)$ = jarak data ke-x ke pusat cluster y

x_i = data ke-i pada atribut data ke-n

y_i = data ke-j pada atribut data ke-n

2. Materials and Methods

Berdasarkan penjelasan sebelumnya, penelitian ini dilakukan melalui beberapa tahapan yang dapat dilihat dalam kerangka tahapan penelitian dibawah ini.



Gambar 1. Tahapan Penelitian

a. Identifikasi Masalah

Langkah pertama dari penelitian ini adalah menggunakan algoritma K-Means untuk mengidentifikasi masalah, dengan tujuan mengetahui jumlah kelompok/kategori yang ada di antara tamu asing di hotel bintang

b. Pengumpulan Data

Proses pengumpulan data yang digunakan berupa studi literatur (Dacwanda & Nataliani, 2021). Dataset untuk penelitian ini diambil dari situs BPS dengan judul Jumlah Tamu Asing Pada Hotel Berbintang pada Tiap Propinsi di Indonesia mulai tahun 2016-2021.

c. Pengolahan Data Awal

Pengolahan data diawali dengan pemilihan table terlebih dahulu kemudian dilakukan pemilihan atribut yang akan digunakan untuk clustering.

d. Implementasi Algoritma K-Means

Jumlah kluster yang digunakan dengan algoritma K-Means yakni sebanyak 3 kluster.

e. Evaluasi Hasil

Dalam perhitungan dengan menggunakan algoritma K-Means, apakah nantinya akan terjadi iterasi dalam pengelompokan datanya. Apabila masih terdapat perubahan maka akan terus dilakukan sampai dengan mendapatkan hasil jarak terdekat.

melalui laman Badan Pusat Statistik (<https://bps.go.id>) yang berjumlah 34 propinsi dari tahun 2016 sampai dengan tahun 2021.

3. Hasil dan Pembahasan

3.1. Analisis Data

Studi ini menggunakan data yang diperoleh

Tabel 1. Data Tamu Asing Menurut Propinsi

Provinsi	Jumlah Tamu Asing pada Hotel Bintang (Ribuan Orang)					
	2016	2017	2018	2019	2020	2021
Aceh	24.34	24.59	11.73	9.21	2.49	0.12
Sumatera utara	311.48	314.78	346.94	201.88	33.25	4.93
Sumatera barat	81.27	75.86	43.08	51.83	10.32	2.37
Riau	33.13	35.19	30.54	35.33	7.77	5.41
Jambi	4.05	7.89	5.66	6.02	1.44	0.81
Sumatera selatan	12.42	14.27	16.96	14.05	2.41	3.23
Bengkulu	1.42	1.07	0.70	0.56	0.09	0.03
Lampung	11.27	9.00	3.25	3.01	0.76	0.57
Kep. Bangka Belitung	10.35	13.17	8.18	9.68	2.19	0.58
Kep. Riau	1294.27	1631.31	2497.80	1595.59	230.52	30.3
Dki Jakarta	1862.32	1879.44	2266.08	1529.76	267.78	283.01
Jawa barat	507.08	516.94	536.60	502.72	118.56	56.81
Jawa tengah	107.17	140.70	112.21	122.15	27.28	15.12
Di Yogyakarta	313.12	317.93	237.59	211.5	37.76	14.2
Jawa timur	269.88	287.27	238.70	267.97	53.95	26.88
Banten	287.72	437.11	496.89	464.21	99.4	66.48
Bali	5701.70	5973.98	9252.94	5687.8	1236.9	97.83

Nusa tenggara barat	285.95	317.10	211.60	195.8	40.62	7.57
Nusa tenggara timur	48.27	49.79	34.78	58.59	11.1	6.1
Kalimantan barat	31.36	35.56	33.33	33	9.21	2.36
Kalimantan tengah	9.31	7.69	5.47	6.68	1.3	1.12
Kalimantan selatan	13.04	13.10	12.10	9.41	2.12	1.05
Kalimantan timur	38.33	37.62	23.41	30.66	10.52	6.5
Kalimantan utara	1.89	2.86	2.44	2.57	1.38	0.43
Sulawesi utara	61.77	128.53	148.36	164.71	12.37	10.19
Sulawesi tengah	3.80	3.86	4.07	2	0.59	0.3
Sulawesi selatan	106.82	109.07	78.22	46.91	14.97	2.13
Sulawesi tenggara	3.96	3.21	2.36	1.4	0.31	0.34
Gorontalo	3.84	5.24	2.94	1.99	0.4	0.3
Sulawesi barat	0.64	0.87	0.19	0.38	0.03	0.01
Maluku	6.17	6.45	10.68	7.57	1.09	0.81
Maluku utara	1.12	0.53	0.63	0.66	0.14	0.13
Papua barat	4.58	7.38	12.60	19.25	4.63	0.97
Papua	12.87	10.89	11.16	12.58	3.89	0.15

3.2. Pengolahan Data

Berikut merupakan tahapan perhitungan dalam melakukan cluster terhadap jumlah tamu asing pada hotel berbintang di Indonesia menurut Propinsi dengan algoritma K-Means:

- Menentukan nilai k, untuk mengetahui jumlah cluster yang akan diinginkan. Dalam penelitian ini, penulis menentukan nilai k=3 untuk jumlah kelompok cluster (cluster 0 = jumlah tamu asing tinggi, cluster 1 = jumlah tamu asing sedang dan cluster 2 = jumlah tamu asing rendah).
- Memilih sejumlah data secara random sebagai centorid atau pusat data. Dalam

penelitian ini, penulis menggunakan data ke- 2 (Propinsi Sumatra Utara), data ke-12 (propinsi Jawa Barat) dan data ke-27 (Propinsi Sulawesi Selatan).

- Menghitung jarak pusat cluster. Proses clusterisasi menggunakan jarak terdekat dari masing-masing data yang dihitung. Dibawah ini adalah penentuan data yang digunakan untuk centroid dan hasil perhitungan data dalam menentukan jarak setiap data dengan masing-masing centroid:

Tabel 2. Penentuan centroid awal

Data	Centroid Data Awal	2016	2017	2018	2019	2020	2021
2 C0		311.48	314.78	346.94	201.88	33.25	4.93
12 C1		507.08	516.94	536.6	502.72	118.56	56.81
27 C2		106.82	109.07	78.22	46.91	14.97	2.13

Proses perhitungan jarak pusat kluster untuk C0, C1 dan C2 menggunakan persamaan 2) yaitu sebagai berikut:

- Aceh

C0

$$\sqrt{(24.34 - 311.48)^2 + (24.59 - 314.78)^2 + (11.73 - 346.94)^2 + (9.21 - 201.88)^2 + (2.49 - 33.25)^2 + (0.12 - 4.93)^2}$$

$$= 234954.163$$

$$\begin{aligned}
 &C1 = \sqrt{(24.34 - 507.08)^2 + (24.59 - 516.94)^2 + (11.73 - 536.6)^2 + (9.21 - 502.72)^2 + (2.49 - 118.56)^2 + (0.12 - 56.81)^2} = 778617.901 \\
 &C2 = \sqrt{(24.34 - 106.82)^2 + (24.59 - 109.07)^2 + (11.73 - 78.22)^2 + (9.21 - 46.91)^2 + (2.49 - 14.97)^2 + (0.12 - 2.13)^2} = 13221.351 \\
 2) \text{ Sumatra Utara} \\
 &C0 = \sqrt{(311.48 - 311.48)^2 + (314.78 - 314.78)^2 + (346.94 - 346.94)^2 + (201.88 - 201.88)^2 + (33.25 - 33.25)^2 + (4.93 - 4.93)^2} = 0 \\
 &C1 = \sqrt{(311.48 - 507.08)^2 + (314.78 - 516.94)^2 + (346.94 - 536.6)^2 + (201.88 - 502.72)^2 + (33.25 - 118.56)^2 + (4.93 - 56.81)^2} = 177509.217 \\
 &C2 = \sqrt{(311.48 - 106.82)^2 + (314.78 - 109.07)^2 + (346.94 - 78.22)^2 + (201.88 - 46.91)^2 + (33.25 - 14.97)^2 + (4.93 - 2.13)^2} = 139089.402 \\
 3) \text{ Sumatra Barat} \\
 &C0 = \sqrt{(81.27 - 311.48)^2 + (75.86 - 314.78)^2 + (43.08 - 346.94)^2 + (51.83 - 201.88)^2 + (10.32 - 33.25)^2 + (2.37 - 4.93)^2} = 172691.217 \\
 &C1 = \sqrt{(81.27 - 507.08)^2 + (75.86 - 516.94)^2 + (43.08 - 536.6)^2 + (51.83 - 502.72)^2 + (10.32 - 118.56)^2 + (2.37 - 56.81)^2} = 656520.77 \\
 &C2 = \sqrt{(81.27 - 106.82)^2 + (75.86 - 109.07)^2 + (43.08 - 78.22)^2 + (51.83 - 46.91)^2 + (10.32 - 14.97)^2 + (2.37 - 2.13)^2} = 2409.1602
 \end{aligned}$$

Sampai dengan perhitungan pada Propinsi Papua
Berikut hasil perhitungan jarak pusat cluster pada iterasi 1

Tabel 3. Jarak masing-masing data dengan centroid

Provinsi	C0	C1	C2	Jarak Terpendek
Aceh	234954.1628	778617.9005	13221.351	13221.351
Sumatera Utara	0	177509.2173	139089.4018	0
Sumatera Barat	172691.217	656520.7701	2409.1602	2409.1602
Riau	206946.2414	722023.5323	8002.0216	8002.0216
Jambi	250350.9306	805096.2005	17461.9114	17461.9114
Sumatera Selatan	235726.225	778357.5263	14072.9708	14072.9708
Bengkulu	260259.4997	823240.1018	20152.8973	20152.8973
Lampung	252548.3211	809710.9378	17861.6235	17861.6235
Kep. Bangka Belitung	243952.5458	793300.8707	15750.6854	15750.6854
Kep. Riau	8342419.424	6296515.786	10618434.54	6296515.786
DKI Jakarta	8028408.197	5977106.77	10264347.41	5977106.77
Jawa Barat	177509.2173	0	598353.9679	0
Jawa Tengah	92102.6792	476966.9391	8137.4608	8137.4608
...

...
...
Maluku Utara	260602.8391	823749.2514	20269.8311	20269.8311
Papua Barat	240773.3985	784571.285	15622.4173	15622.4173
Papua	242115.2985	789264.7902	15535.5417	15535.5417

d. Mengelompokkan objek berdasarkan jarak terdekat.

Apabila nilai tertinggi masuk ke dalam cluster 0, maka data masuk ke dalam cluster 0, apabila nilai terendah maka masukkan ke dalam kelompok cluster 2, nilai sedang maka masukkan data ke dalam kelompok cluster 1. Hasil pengelompokkan cluster ditunjukkan pada tabel 3 berikut:

Papua Barat			1
Papua			1

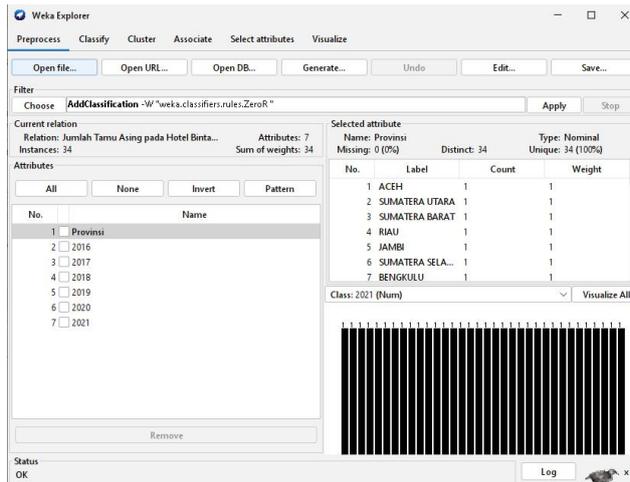
Tabel 4. Data Hasil Clusterisasi

Provinsi	C0	C1	C2
Aceh			1
Sumatera Utara	1		
Sumatera Barat			1
Riau			1
Jambi			1
Sumatera Selatan			1
Bengkulu			1
Lampung			1
Kep. Bangka Belitung			1
Kep. Riau		1	
DKI Jakarta		1	
Jawa Barat		1	
Jawa Tengah			1
DI Yogyakarta	1		
Jawa Timur	1		
Banten		1	
Bali		1	
Nusa Tenggara Barat	1		
Sulawesi Tengah			1
Sulawesi Selatan			1
Sulawesi Tenggara			1
...
...
...
Maluku Utara			1

Proses perhitungan dengan algoritma K-Means akan melakukan proses iterasi ke-n sampai dengan data pengelompokkan tidak berubah. Setelah menentukan nilai centroid untuk iterasi berikutnya, maka dilakukan perhitungan kembali untuk mendapatkan jarak terdekat. Pada penelitian ini, proses iterasi dilakukan sampai dengan 3 iterasi. Berdasarkan tabel 3 diatas, hasil menunjukkan bahwa cluster 0 (4 propinsi), cluster 1 (6 propinsi) dan cluster 2 (24 propinsi).

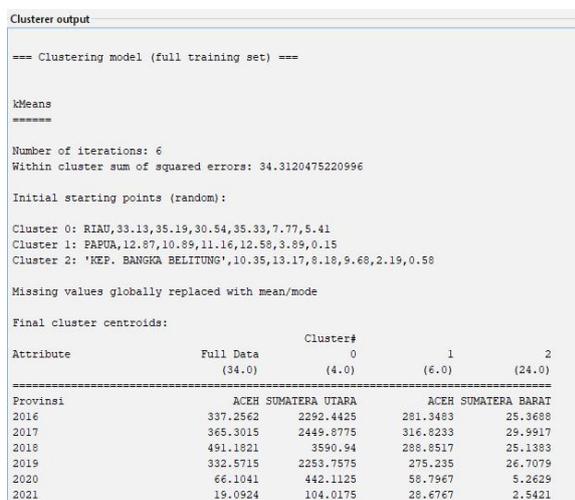
3.3. Implementasi WEKA

Pengolahan data yang dilakukan dalam penelitian menggunakan aplikasi Weka 3.8.6. Proses pertama yang dilakukan adalah mengaplikasikan data yang digunakan, yakni data tamu asing pada hotel bintang menurut propinsi seperti pada gambar berikut:



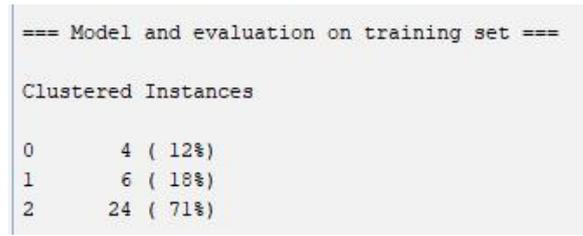
Gambar 2. Aplikasi Dataset

Setelah melakukan aplikasi data pada Weka, langkah selanjutnya adalah melakukan pemodelan dengan melakukan proses clusterisasi dengan algoritma K-Means seperti pada gambar berikut:

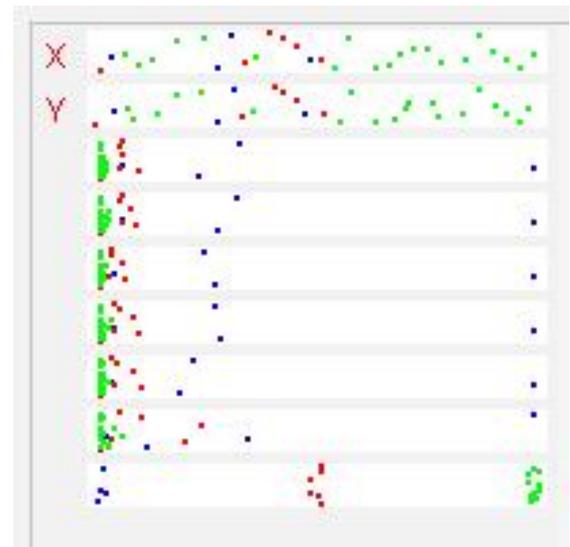


Gambar 3. Cluster Output

Data penelitian terdiri dari 34 data dengan 7 atribut. Dalam proses perhitungan cluster, dibuat 3 cluster yakni cluster 0 (tinggi), cluster 1 (sedang) dan cluster 2 (rendah).



Gambar 4. Kelompok Cluster



Gambar 5. Visualisasi Cluster pada Weka

4. Kesimpulan

Berdasarkan perhitungan yang dilakukan secara manual dan dengan menggunakan aplikasi Weka 3.8.6 dan perhitungan dengan algoritma K-Means untuk clustering tamu asing pada hotel berbintang menurut propinsi menghasilkan 3 cluster yaitu cluster 0 (tinggi) ada 4 propinsi, cluster 1 (sedang) sebanyak 6 propinsi dan cluster 2 (rendah) sebanyak 24 propinsi.

Dengan adanya pengelompokan data tersebut propinsi dengan cluster 0 memiliki daya tarik wisata yang sangat tinggi sehingga jumlah tamu asing yang menginap di hotel bintang pun lebih banyak

dibandingkan propinsi dengan cluster 1 dan cluster 2.

Untuk penelitian selanjutnya dapat dilakukan dengan penggunaan tools yang lain seperti Rapid Miner, Orange dan yang lainnya serta penambahan data yang dijadikan objek penelitian sehingga data yang didapatkan lebih bervariasi lagi.

Acknowledgments

Penulis memahami bahwa menulis artikel ini tidaklah mudah. Oleh karena itu, penulis mengucapkan terima kasih kepada semua pihak yang telah mendukung penulisan artikel ini. Semoga artikel ini dapat dikembangkan lebih lanjut di masa mendatang, terutama dalam proses implementasi algoritma.

References

- Atma, Y. A., Montesori, S., Teknologi, J., Politeknik, I., & Padang, N. (2022). *ANALISIS DATA MINING UNTUK MENENTUKAN PROFIT PERUSAHAAN MENGGUNAKAN METODE K-MEANS THE DATA MINING ANALYSIS TO DETERMINE COMPANY PROFIT USING K-MEANS METHOD*. 3(2), 29–36.
- Dacwanda, D. O., & Nataliani, Y. (2021). Implementasi k-Means Clustering untuk Analisis Nilai Akademik Siswa

Berdasarkan Nilai Pengetahuan dan Keterampilan. *Aiti*, 18(2), 125–138. <https://doi.org/10.24246/aiti.v18i2.125-138>

Direktur Jenderal Pariwisata Nomor : 14/U/II/88. (1988). Keputusan Direktur Jenderal Pariwisata No. 14/U/II/88. *Pedoman Pembinaan Pramuwisata Khusus Direktorat Jenderal Pariwisata*, 362. https://jdih.kemendparekraf.go.id/asset/data_puu/regulation_subject_1576230513_14uui88.pdf

Indonesia, P. R. (2007). *Peraturan Presiden No. 86 Tahun 2007 tentang Badan Pusat Statistik*. http://www2.pom.go.id/public/hukum_perundangan/pdf/Pengamanan_rokok_bagi_kesehatan.pdf

Khotimah, T. (2014). Pengelompokan Surat Dalam Al Qur'an Menggunakan Algoritma K-Means. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 5(1), 83–88. <https://doi.org/10.24176/simet.v5i1.141>

Kurniawan, W., Rifai, A., Gata, ; Windu, & Gunawan, D. (2020). Analisis Algoritma K-Medoids Clustering Dalam Menentukan Pemesanan Hotel. *Jurnal Swabumi*, 8(2), 182–187.

- <https://www.kaggle.com>
- Listiani, L., Agustin, Y. H., & Ramdhani, M. Z. (2019). Implementasi algoritma k-means cluster untuk rekomendasi pekerjaan berdasarkan pengelompokan data penduduk. *SENSITif: Seminar Nasional Sistem Informasi Dan Teknologi Informasi*, 761–769.
- Nasir, J. (2021). Penerapan Data Mining Clustering Dalam Mengelompokan Buku Dengan Metode K-Means. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 11(2), 690–703. <https://doi.org/10.24176/simet.v11i2.5482>
- Pramayuda, D., Lubis, M. R., Saragih, I. S., & Tunas Bangsa, S. (2020). Pengelompokan Tamu Asing Ke Indonesia Berdasarkan Provinsi Dengan Algoritma K-Means. *Kesatria : Jurnal Penerapan Sistem Informasi (Komputer Dan Manajemen)*, 1(4), 130–137. <http://tunasbangsa.ac.id/pkm/index.php/kesatria/article/view/38>
- Purba, W., Siawin, W., & . H. (2019). Implementasi Data Mining Untuk Pengelompokan Dan Prediksi Karyawan Yang Berpotensi Phk Dengan Algoritma K-Means Clustering. *Jurnal Sistem Informasi Dan Ilmu Komputer Prima(JUSIKOM PRIMA)*, 2(2), 85–90. <https://doi.org/10.34012/jusikom.v2i2.429>
- Sari, R. W., & Hartama, D. (2018). Data Mining : Algoritma K-Means Pada Pengelompokan Wisata Asing ke Indonesia Menurut Provinsi. *Seminar Nasional Sains & Teknologi Informasi (SENSASI)*, 322–326.