

Pengelompokan Tamu Asing Ke Indonesia Berdasarkan Provinsi Dengan Algoritma K-Means

Dicki Pramayuda¹, Muhammad Ridwan Lubis², Ilham Syahputra Saragih³

^{1,3}STIKOM Tunas Bangsa, Pematangsiantar, Medan, Indonesia

²AMIK Tunas Bangsa, Pematangsiantar, Medan, Indonesia

¹dickipramayuda@gmail.com, ²ridwanlubis@amiktunasbangsa.ac.id,

³ilham@amiktunasbangsa.ac.id

Abstract

The development of the industry in Indonesia has been growing rapidly, this is marked by the number of foreign guests who come to Indonesia to make a visit as an object of tourist attraction. In addition, which makes foreign guests feel happy coming to Indonesia, the friendliness of the population is also one of the factors that are very supportive. Lack of information and grouping regarding the visit of foreign guests to various tourist attraction objects in various provinces resulted in difficulties for the Regional Government, especially the Office of Tourism and Creative Economy in planning tourism marketing strategies in various provinces so that tourism promotion or marketing is not on target, causing tourism- tours in lesser-known provinces are rarely visited by foreign guests. The goal to be achieved in this study is to get a grouping of data on foreign guest visits in each province during the period 2014-2018 using the K-Means algorithm. K-means is a non-hierarchical data clustering method that attempts to partition existing data into one or more clusters / groups. So that the results obtained can be considered by the government, especially the Office of Tourism and Creative Economy in the Provinces to develop the potential of tourist attraction objects in the provinces in Indonesia.

Keywords : Data Mining, Foreign Guests, Clustering, K-Means

Abstrak

Perkembangan industri di Indonesia sudah berkembang dengan pesat, hal ini ditandai dengan banyaknya tamu asing yang datang ke Indonesia untuk melakukan kunjungan sebagai objek daya tarik wisata. Selain itu, yang membuat para tamu asing merasa senang datang ke Indonesia, keramah-tamahan penduduknya juga merupakan salah satu faktor yang sangat mendukung. Kurangnya informasi serta pengelompokan menyangkut kunjungan tamu asing ke berbagai obyek daya tarik wisata di berbagai provinsi mengakibatkan adanya kesulitan bagi Pemerintah Daerah khususnya Dinas Pariwisata dan Ekonomi Kreatif dalam perencanaan strategi pemasaran pariwisata di berbagai provinsi sehingga promosi atau pemasaran pariwisata menjadi tidak tepat sasaran, menyebabkan wisata-wisata di provinsi-provinsi yang kurang dikenal jarang di kunjungi oleh tamu asing. Adapun tujuan yang ingin dicapai dalam penelitian ini adalah mendapatkan pengelompokan dari data kunjungan tamu asing pada setiap provinsi-provinsi selama kurun waktu tahun 2014-2018 menggunakan Algoritma K-Means. K-means merupakan salah satu metode data clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster / kelompok. Sehingga hasil yang diperoleh dapat menjadi pertimbangan pemerintah khususnya Dinas Pariwisata dan Ekonomi Kreatif di Provinsi-provinsi untuk mengembangkan potensi objek daya tarik wisata di provinsi yang ada di Indonesia.

Kata Kunci : Data Mining, Tamu Asing, Clustering, K-Means

1. Pendahuluan

Indonesia merupakan salah satu negara yang kaya akan keindahan wisata alam, taman wisata dan taman budaya yang dapat mendukung perkembangan pariwisata di Indonesia. Kondisi letak geografis Indonesia yang kaya akan keindahan alam menjadikan tamu asing tertarik untuk mengunjungi Indonesia untuk melihat keindahan alam tersebut. Maka dengan pengelolaan yang baik dan terarah diharapkan pariwisata di Indonesia mampu menarik tamu asing untuk berkunjung. Hal itu menjadikan pendapatan negara bertambah dan juga besarnya minat wisatawan berkunjung ke Indonesia memicu meningkatkan peluang dalam jenis berbisnis baik objek wisatanya itu sendiri maupun usaha-usaha lain yang berhubungan dengan aktivitas wisata seperti akomodasi hotel, transportasi, kuliner dan lain sebagainya.

Sumber data dari penelitian ini dikumpulkan berdasarkan dari Jumlah Tamu Asing ke Indonesia berdasarkan Provinsi yang dihasilkan oleh Badan Pusat Statistik Nasional. Badan Pusat Statistik adalah Lembaga Pemerintah Non Kementerian yang bertanggung jawab langsung kepada Presiden. Sebelumnya, Badan Pusat Statistik merupakan Biro Pusat Statistik, yang dibentuk berdasarkan UU Nomor 6 Tahun 1960 tentang Sensus dan UU Nomor 7 Tahun 1960 tentang Statistik. Diharapkan Penelitian ini dapat menjadi upaya dalam membantu mengembangkan potensi objek daya tarik wisata yang dilakukan dengan mengelompokkan data kunjungan tamu asing pada setiap provinsi yang ada di Indonesia.

2. Metodologi Penelitian

2.1. K-Means

K-Means clustering merupakan salah satu metode data *clustering non-hirarki* yang mengelompokkan data dalam bentuk satu atau lebih *cluster*/kelompok [1]. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster* kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan *cluster*/kelompok yang lain sehingga data yang berada dalam satu *cluster*/kelompok memiliki tingkat variasi yang kecil [2].

K-Means merupakan suatu algoritma pengklasteran yang cukup sederhana yang mempartisi dataset kedalam beberapa klaster *k*. Algoritmanya cukup mudah untuk diimplementasi dan dijalankan, relatif cepat, mudah disesuaikan dan banyak digunakan menurut Wu & Kumar [3]. Tahapan melakukan *clustering* atau pengelompokan dengan metode *K-Means* adalah sebagai berikut [4].

- Menentukan berapa banyak *cluster* yang ingin yang ingin dibentuk, di mana nilai *K* adalah banyaknya *cluster*/ jumlah *cluster*.
- Menentukan pusat *cluster* (*centroid*) awal.
- Setelah menentukan *centroid* awal, maka setiap data akan menemukan *centroid* terdekatnya.
- Setelah menghitung jarak data ke *centroid*nya.
- Mencari *centroid* baru berdasarkan membership dari masing-masing *cluster*.
- Kembali ke tahap 3.
- Perulangan berhenti apabila tidak ada data yang berpindah.

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

Di mana :

v_{ij} = *centroid* rata-rata *cluster* ke-i untuk variable k-j

N_i = jumlah anggota *cluster* ke-i

i, k = indeks dari *cluster*

j = indeks dari variable

X_{ki} = nilai data ke-k variable ke-j dalam *cluster* tersebut.

Menggunakan rumus *Euclidean Distance* berikut:

$$d_{Euclidean}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

Di mana :

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

x_i = data ke- i pada atribut data ke n

y_j = data ke- j pada atribut data ke n .

2.2. Tamu Asing atau Wisatawan

Menurut undang – undang No. 10 tahun 2009 tentang kepariwisataan, disebutkan tamu asing atau wisatawan adalah orang yang melakukan wisata. Sedangkan wisata adalah bepergian secara bersama-sama dengan tujuan untuk bersenang-senang, menambah pengetahuan dan lain lain. Sehingga dapat diartikan tamu asing atau wisatawan adalah orang-orang yang datang berkunjung pada suatu tempat atau negara yang bukan tempat tinggalnya untuk tujuan tertentu [5].

3. Hasil dan Pembahasan

3.1. Analisa Data

Penelitian ini dilakukan berdasarkan data yang diperoleh dari Badan Pusat Statistik (BPS) yang terdiri dari 34 provinsi yang dijadikan sebagai alternatif dengan 5 kriteria yaitu mulai tahun 2014 sampai dengan 2018.

Tabel 1. Data Jumlah Tamu Asing Berdasarkan Provinsi

Provinsi	Jumlah Tamu Asing Berdasarkan Provinsi (Ribu Orang)				
	2014	2015	2016	2017	2018
Aceh	11.80	13	24.34	24.59	11.73
Sumatera Utara	453.20	263.10	311.48	314.78	346.94
Sumatera Barat	63.80	38	81.27	75.86	43.08
Riau	88.70	25.90	33.13	35.19	30.54
Jambi	2.10	4.90	4.05	7.89	5.66
Sumatera Selatan	48.40	13.20	12.42	14.27	16.96
Bengkulu	0.80	0.60	1.42	1.07	0.70
Lampung	17.90	4.60	11.27	9	3.25
Kep. Bangka Belitung	2.60	3.30	10.35	13.17	8.18
Kep. Riau	1439.70	1329	1294.27	1631.31	2497.80
Dki Jakarta	1222.70	1873.80	1862.32	1879.44	2266.08
Jawa Barat	333	629.30	507.08	516.94	536.60
Jawa Tengah	120.90	104.90	107.17	140.70	112.21
Di Yogyakarta	202.80	190.10	313.12	317.93	237.59
Jawa Timur	232	249.90	269.88	287.27	238.70
Banten	174.50	351.20	287.72	437.11	496.89
Bali	5293.50	5337.80	5701.70	5973.98	9252.94
Nusa Tenggara Barat	220.40	154.20	285.95	317.10	211.60
Nusa Tenggara Timur	22.60	25	48.27	49.79	34.78
Kalimantan Barat	18.40	24.90	31.36	35.56	33.33
Kalimantan Tengah	3.20	4.10	9.31	7.69	5.47
Kalimantan Selatan	8.30	11.60	13.04	13.10	12.10
Kalimantan Timur	52.40	39.20	38.33	37.62	23.41
Kalimantan Utara	-	1.30	1.89	2.86	2.44
Sulawesi Utara	43.70	12.50	61.77	128.53	148.36
Sulawesi Tengah	1.30	1.80	3.80	3.86	4.07
Sulawesi Selatan	66.10	66.90	106.82	109.07	78.22
Sulawesi Tenggara	5.80	5.10	3.96	3.21	2.36
Gorontalo	0.40	0.50	3.84	5.24	2.94
Sulawesi Barat	-	0.10	0.64	0.87	0.19
Maluku	16.50	66.90	6.17	6.45	10.68
Maluku Utara	0.30	0.30	1.12	0.53	0.63
Papua Barat	2.70	4.80	4.58	7.38	12.60
Papua	14.20	17.90	12.87	10.89	11.16

3.2. Pengolahan Data

Berikut ini merupakan langkah-langkah yang dilakukan penulis dalam mengelompokkan Jumlah Tamu Asing yang datang ke Indonesia berdasarkan Provinsi menggunakan Algoritma *K-Means*:

a) Menentukan data yang ingin di *cluster*

Data yang digunakan pada proses clustering yaitu Data Jumlah Tamu Asing yang datang ke Indonesia berdasarkan Provinsi mulai tahun 2014 sampai dengan 2018 dengan menggunakan data sebanyak 34 Provinsi yang ada di Indonesia. Pada proses *clustering* diawali dengan mencari nilai rata-rata untuk setiap Provinsi.

Berikut ini beberapa contoh mencari nilai rata – rata :

$$R1 = 11,8 + 13 + 24,34 + 24,59 + 11,73 / 5 = 17,09$$

$$R2 = 453,2 + 263,1 + 311,48 + 314,78 + 346,94 / 5 = 337,90$$

$$R3 = 63,8 + 38 + 81,27 + 75,86 + 43,08 / 5 = 60,40$$

Perhitungan mencari nilai rata-rata dilanjutkan sampai dengan R34. Dalam melakukan *clustering*, data yang diperoleh akan dihitung terlebih dahulu berdasarkan hasil rata-rata Jumlah Tamu Asing yang datang ke Indonesia berdasarkan Provinsi.

Data tersebut kemudian diakumulasikan dan diambil nilai rata-rata nya, yaitu :

Tabel 2. Jumlah Tamu Asing berdasarkan Provinsi

Provinsi	Rata-Rata	Provinsi	Rata-Rata
Aceh	17,09	Nusa Tenggara Barat	237,85
Sumatera Utara	337,90	Nusa Tenggara Timur	36,09
Sumatera Barat	60,40	Kalimantan Barat	28,71
R i a u	42,69	Kalimantan Tengah	5,95
J a m b i	4,92	Kalimantan Selatan	11,63
Sumatera Selatan	21,05	Kalimantan Timur	38,19
Bengkulu	0,92	Kalimantan Utara	2,12
Lampung	9,20	Sulawesi Utara	78,97
Kep Bangka Belitung	7,52	Sulawesi Tengah	2,97
Kepulauan Riau	1638,42	Sulawesi Selatan	85,42
DKI Jakarta	1820,87	Sulawesi Tenggara	4,09
Jawa Barat	504,58	Gorontalo	2,58
Jawa Tengah	117,18	Sulawesi Barat	0,45
DI Yogyakarta	252,31	M a l u k u	21,34
Jawa Timur	255,55	Maluku Utara	0,58
Banten	349,48	Papua Barat	6,41
B a l i	6311,98	P a p u a	13,40

b) Menentukan *centroid* data awal

Dalam penerapan algoritma *K-Means* dihasilkan nilai titik tengah atau *centroid* dari data yang didapat dengan ketentuan bahwa *clusterisasi* yang diinginkan adalah 2. Penentuan titik *cluster* ini dilakukan dengan mengambil nilai terbesar (*maksimum*) untuk *cluster* tinggi (C1) dan nilai terkecil (*minimum*) untuk *cluster* rendah (C2). Berikut ini merupakan hasil dari *centroid* data awal :

Tabel 3. *Centroid* Data Awal

C1	6311,98
C2	0,45

c) Menghitung jarak pusat *cluster*

Dengan menggunakan *centroid* tersebut maka dapat di *cluster* data yang telah didapat menjadi 2 *cluster*. Proses *cluster* dengan mengambil jarak terdekat dari setiap data yang diolah. Berikut ini contoh perhitungan jarak setiap data pada *cluster* pertama (C1) :

$$D(1,1) = \sqrt{(6311,98 - 17,09)^2} = 6294,89$$

$$D(1,2) = \sqrt{(6311,98 - 337,90)^2} = 5974,08$$

$$D(1,3) = \sqrt{(6311,98 - 60,40)^2} = 6251,58$$

Sampai dengan D (1,34)

Perhitungan jarak pada setiap data pada cluster kedua(C2) :

$$D(2,1) = \sqrt{(0,45 - 17,09)^2} = 16,64$$

$$D(2,2) = \sqrt{(0,45 - 337,90)^2} = 337,45$$

$$D(2,3) = \sqrt{(0,45 - 60,40)^2} = 59,95$$

Sampai dengan D (2,34)

Berikut tabel 4. hasil perhitungan jarak data dengan titik pusat pada Literasi 1:

Tabel 4. Perhitungan Jarak Pusat *Cluster* Literasi 1

Provinsi	Rata-Rata	Literasi 1		
		C1	C2	Jarak Terpendek
Aceh	17,09	6294,89	16,64	16,64
Sumatera Utara	337,90	5974,08	337,45	337,45
Sumatera Barat	60,40	6251,58	59,95	59,95
Riau	42,69	6269,29	42,24	42,24
Jambi	4,92	6307,06	4,47	4,47
Sumatera Selatan	21,05	6290,93	20,60	20,60
Bengkulu	0,92	6311,07	0,47	0,47
Lampung	9,20	6302,78	8,75	8,75
Kep. Bangka Belitung	7,52	6304,46	7,07	7,07
Kep. Riau	1638,42	4673,57	1637,97	1637,97
Dki Jakarta	1820,87	4491,12	1820,42	1820,42
...
...
Maluku	21,34	6290,64	20,89	20,89
Maluku Utara	0,58	6311,41	0,13	0,13
Papua Barat	6,41	6305,57	5,96	5,96
Papua	13,40	6298,58	12,95	12,95

d) Mengelompokkan setiap data berdasarkan kedekatannya dengan *centroid* (jarak terkecil).

Jika nilai tertinggi terdapat di *cluster* 1 (C1) maka masuk kedalam kelompok *cluster* 1 dan begitu juga sebaliknya, dapat dilihat pada tabel 5. sebagai berikut:

Tabel 5. Hasil *Cluster* Literasi 1

Provinsi	C1	C2	Provinsi	C1	C2
Aceh		1	Nusa Tenggara Barat		1
Sumatera Utara		1	Nusa Tenggara Timur		1
Sumatera Barat		1	Kalimantan Barat		1
R i a u		1	Kalimantan Tengah		1
J a m b i		1	Kalimantan Selatan		1
Sumatera Selatan		1	Kalimantan Timur		1
Bengkulu		1	Kalimantan Utara		1
Lampung		1	Sulawesi Utara		1
Kep Bangka Belitung		1	Sulawesi Tengah		1
Kepulauan Riau		1	Sulawesi Selatan		1
DKI Jakarta		1	Sulawesi Tenggara		1
Jawa Barat		1	Gorontalo		1
Jawa Tengah		1	Sulawesi Barat		1
DI Yogyakarta		1	M a l u k u		1
Jawa Timur		1	Maluku Utara		1
Banten		1	Papua Barat		1
B a l i	1		P a p u a		1

Proses *K-Means* akan terus berliterasi sampai pengelompokan data sama dengan pengelompokan data literasi sebelumnya. Dengan kata lain, proses akan terus melakukan

literasi sampai data pada literasi terakhir sama dengan literasi sebelumnya. Setelah mendapatkan nilai titik tengah atau *centroid*, proses sama dilakukan dengan mencari jarak terdekat. Berikut ini contoh perhitungan jarak setiap data pada *cluster* pertama (C1) pada literasi 2 :

$$C1 = \frac{6311,98}{1} = 6312$$

$$C2 = \frac{17,09+337,90+60,40+42,69+4,92+21,05+0,92+9,20+7,52+1638,42+1820,87+504,58+117,18+252,31+255,55+349,48+237,85+36,09+28,71+5,95+11,63+38,19+2,12+78,97+2,97+85,42+4,09+2,58+0,45+21,34+0,58+6,41+13,40}{33} = 182,329$$

Tabel 6. Hasil Perhitungan *Centroid* baru pada Literasi 2

C1	6312
C2	182,329

Setelah mendapatkan titik pusat *cluster* (*centroid*) baru maka langkah selanjutnya yaitu menghitung kembali jarak setiap data jumlah tamu asing berdasarkan provinsi terhadap masing-masing pusat *cluster*. Perhitungan jarak setiap data pada *cluster* pertama (C1) :

$$E(1,1) = \sqrt{(6312 - 17,09)^2} = 6294,89$$

$$E(1,2) = \sqrt{(6312 - 337,90)^2} = 5974,08$$

$$E(1,3) = \sqrt{(6312 - 60,40)^2} = 6251,58$$

Sampai dengan E (1,34)

Perhitungan jarak pada setiap data pada *cluster* kedua (C2) :

$$E(2,1) = \sqrt{(182,33 - 17,09)^2} = 165,24$$

$$E(2,2) = \sqrt{(182,33 - 337,90)^2} = 155,57$$

$$E(2,3) = \sqrt{(182,33 - 60,40)^2} = 121,93$$

Sampai dengan E (2,34)

Berikut hasil perhitungan jarak data dengan titik pusat pada literasi 2.

Tabel 7. Perhitungan Jarak Pusat *Cluster* Literasi 2

Provinsi	Rata-Rata	Literasi 2		
		C1	C2	Jarak Terpendek
Aceh	17,09	6294,89	165,24	165,24
Sumatera Utara	337,9	5974,08	155,57	155,57
Sumatera Barat	60,4	6251,58	121,93	121,93
Riau	42,69	6269,29	139,64	139,64
Jambi	4,92	6307,06	177,41	177,41
Sumatera Selatan	21,05	6290,93	161,28	161,28
Bengkulu	0,92	6311,07	181,41	181,41
Lampung	9,2	6302,78	173,12	173,12
Kep. Bangka Belitung	7,52	6304,46	174,81	174,81
Kep. Riau	1638,42	4673,57	1456,09	1456,09
Dki Jakarta	1820,87	4491,12	1638,54	1638,54
Jawa Barat	504,58	5807,40	322,26	322,26
...
...
Maluku Utara	0,58	6311,41	181,75	181,75
Papua Barat	6,41	6305,57	175,92	175,92
Papua	13,4	6298,58	168,92	168,92

Berikut hasil perhitungan pada data *cluster* literasi 2 yaitu :

Tabel 8. Hasil *Cluster* Literasi 2

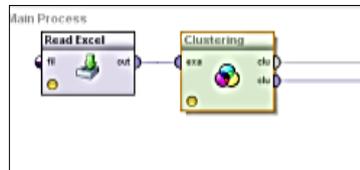
Provinsi	C1	C2	Provinsi	C1	C2
Aceh		1	Nusa Tenggara Barat		1
Sumatera Utara		1	Nusa Tenggara Timur		1

Provinsi	C1	C2	Provinsi	C1	C2
Sumatera Barat		1	Kalimantan Barat		1
R i a u		1	Kalimantan Tengah		1
J a m b i		1	Kalimantan Selatan		1
Sumatera Selatan		1	Kalimantan Timur		1
Bengkulu		1	Kalimantan Utara		1
Lampung		1	Sulawesi Utara		1
Kep Bangka Belitung		1	Sulawesi Tengah		1
Kepulauan Riau		1	Sulawesi Selatan		1
DKI Jakarta		1	Sulawesi Tenggara		1
Jawa Barat		1	Gorontalo		1
Jawa Tengah		1	Sulawesi Barat		1
DI Yogyakarta		1	M a l u k u		1
Jawa Timur		1	Maluku Utara		1
Banten		1	Papua Barat		1
B a l i	1		P a p u a		1

Setelah hasil pada literasi 2 di dapatkan, bandingkan hasil pada literasi 1 dan literasi 2, jika tidak ada perubahan maka proses berhenti, jika ada perubahan maka proses akan dilanjutkan kembali hingga mendapatkan hasil yang sama. Berdasarkan perhitungan manual pada data jumlah tamu asing yang datang ke Indonesia berdasarkan Provinsi yang telah dilakukan diatas mendapatkan hasil akhir literasi 2 bernilai sama yaitu C1= 1 dan C2= 33. Posisi pada literasi tidak berubah, maka proses berhenti dan disimpulkan:

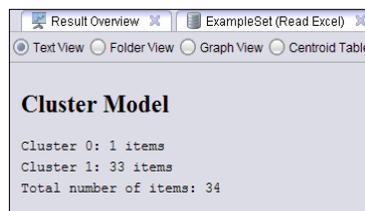
3.3. Implementasi RapidMiner

Data telah berhasil di *import* maka langkah selanjutnya yaitu mengolah data menggunakan *Algoritma K-Means Clustering*. Masukkan Operator *K-Means Clustering* kedalam lembar kerja main proses lalu hubungkan *read excel* dengan operator *K-Means Clustering*, dan tentukan jumlah *cluster*.



Gambar 1. Pemrosesan Data

Setelah semua terkoneksi dan tidak ada yang error kemudian klik tombol *run* pada *tool box* untuk mendapatkan hasil *Clustering* dapat dilihat pada gambar berikut:



Gambar 2. Hasil Cluster Model

Keterangan:

1. *Cluster0* (Tinggi) sebanyak 1 Provinsi
2. *Cluster1* (Rendah) sebanyak 33 Provinsi
3. Jumlah Keseluruhan yaitu 34 Provinsi

4. Kesimpulan

Berdasarkan pembahasan dan hasil *implementasi software RapidMiner* maka diperoleh kesimpulan dalam menentukan Jumlah Tamu Asing yang datang ke Indonesia berdasarkan Provinsi adalah sebagai berikut :

- a) Dari hasil perhitungan dengan menggunakan algoritma *K-Means* dalam pengelompokan Jumlah Tamu Asing yang datang ke Indonesia berdasarkan Provinsi maka Provinsi yang mendapat urutan tertinggi sebanyak 1 Provinsi yaitu Provinsi Bali dan urutan terendah sebanyak 33 Provinsi. Dari hasil perhitungan tersebut maka dapat dijadikan upaya dalam membantu mengembangkan potensi objektif daya tarik wisata yang dilakukan dengan mengelompokkan data kunjungan tamu asing.
- b) Aplikasi yang digunakan yaitu *software RapidMiner5.3* yang dibuat sebagai penyesuaian hasil dari perhitungan manual yang nantinya akan dibandingkan dengan hasil yang ditampilkan *software RapidMiner5.3*.

Daftar Pustaka

- [1] A. P. Windarto, U. Indriani, M. R. Raharjo, and L. S. Dewi, "Bagian 1: Kombinasi Metode Klastering dan Klasifikasi (Kasus Pandemi Covid-19 di Indonesia)," *J. Media Inform. Budidarma*, vol. 4, no. 3, p. 855, 2020, doi: 10.30865/mib.v4i3.2312.
- [2] dan G. F. Ade Bastian, Harun Sujadi, "Penerapan Algoritma K-Means Clustering Analysis Pada Penyakit Menular Manusia (Studi Kasus Kabupaten Majalengka)," no. 1, pp. 26–32, 2018.
- [3] J. S. Novi Yona Sidratul Munti, Gunadi Widi Nurcahyo, "Analisis Dan Penerapan Data Mining Untuk Menentukan Gaji Karyawan Tetap Dan Karyawan Kontrak Menggunakan Algoritma K-Means Clustering (Studi Kasus Di Pt Indomex Dwijaya Lestari)," 2018.
- [4] H. Windania Purba, Willy Siawin, "Implementasi Data Mining Untuk Pengelompokan Dan Prediksi Karyawan Yang Berpotensi Phk Dengan Algoritma K-Means Clustering," vol. 2, no. 2, 2019.
- [5] R. W. Sari and D. Hartama, "Data Mining: Algoritma K-Means Pada Pengelompokan Wisata Asing ke Indonesia Menurut Provinsi," *Semin. Nas. Sains Teknol. Inf.*, pp. 322–326, 2018.