

Implementasi Data Mining Untuk Pengelompokan Data Pengunjung Museum Menggunakan Algoritma K-Means Clustering Pada Museum Deli Serdang

Natasah Liaza*, Yohanni Syahra**, Ita Mariami**

* Sistem Informasi, STMIK Triguna Dharma

** Sistem Informasi, STMIK Triguna Dharma

Article Info

Article history:

Keyword:

Museum Deli Serdang,
Pengelompokan Data
Pengunjung Pada Museum
Data Mining
Algoritma K-Means Clustering

ABSTRACT

Museum Deli serdang merupakan museum sejarah dan budaya yang didirikan dengan tujuan mencatat sejarah Deli Serdang dan memperkenalkannya pada masyarakat .

Meningkatkan jumlah pengunjung museum sangat disarankan bagi pengelola museum, dikarenakan museum merupakan sektor pariwisata yang memiliki daya tarik bagi wisatawan untuk hiburan maupun pembelajaran terhadap benda-benda peninggalan sejarah. Dalam upaya mempertahankan dan meningkatkan jumlah pengunjung museum, melakukan suatu pengelompokan data pengunjung museum sangat disarankan bagi pengelola museum, untuk mengetahui Museum Deli Serdang yang memiliki jumlah pengunjung tertinggi maupun terendah.

Berdasarkan masalah diatas maka diperlukan suatu aplikasi berupa data mining dengan penerapan algoritma K-Means Clustering diharapkan dapat membantu pihak pengurus Museum Deli Serdang dalam meningkatkan jumlah kunjungan pengunjung yang ada pada Museum Deli Serdang.

Hasil dari pengelompokan ini data dikelompokkan menjadi 3 cluster yaitu C1= jumlah kunjungan pengunjung sedang, C2= jumlah kunjungan pengunjung tinggi, C3= jumlah kunjungan pengunjung rendah.

Copyright © 2020 STMIK Triguna Dharma.

All rights reserved.

Corresponding Author:

Nama : Natasah Liaza
Sistem Informasi
STMIK Triguna Dharma
Email: natasahliaza25@gmail.com

1. PENDAHULUAN

Museum merupakan tempat untuk mengoleksi, menyimpan, merawat dan memamerkan benda-benda peninggalan budaya yang berasal dari zaman berbeda. Pengunjung museum memiliki peran yang penting terhadap pengembangan museum mengingat bahwa kunjungan ke museum-museum di Indonesia masih belum banyak dilakukan oleh masyarakat, dan adanya jumlah pengunjung museum dapat digunakan oleh pihak museum untuk mengembangkan museum dari berbagai aspek [1].

Untuk mempermudah pengelola museum dalam mempertahankan dan mengembangkan jumlah pengunjung museum, maka dapat diatasi dengan menggunakan data *mining*. Didalam data *mining* terdapat beberapa metode salah satunya adalah metode pengelompokan algoritma *k-means clustering*. Metode pengelompokan algoritma *k-means clustering* adalah proses pengelompokan sejumlah data/objek ke dalam kelompok data sehingga setiap kelompok berisi data yang mirip.

2. METODE PENELITIAN

2.1 Data Mining

Data *Mining* adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa informasi tersembunyi dalam sekumpulan absis data yang besar. Data *Mining* merupakan kegiatan yang meliputi pengumpulan, pemakaian dan histori untuk menemukan informasi baru dengan pola dan hubungan dalam *set* data berukuran besar. Hasil dari pengolahan data menggunakan metode data *mining* ini dapat bermanfaat untuk mengambil keputusan di masa depan serta mudah di mengerti [2].

Data *mining* merupakan suatu bagian langkah yang penting dalam proses *knowledge discovery in database* terutama berkaitan dengan ekstraksi dan perhitungan data-data yang ditelaah. Adapun proses KDD adalah suatu proses menggunakan *database* dalam jumlah besar untuk mencari dan mengidentifikasi pola (*pattern*) dalam data, dimana pola yang ditemukan bersifat sah, baru dapat bermanfaat dan dapat dimengerti [3]. Tahapan proses dalam penggunaan Data *Mining* yang merupakan proses *Knowledge Discovery In Database* (KDD) yaitu *Data Selection, Pre-processing/Cleaning, Transformation, Data Mining, Interpretation/Evaluation*.

2.2 Algoritma K-Means Clustering

K-Means termasuk dalam *partitioning clustering* yaitu setiap data harus masuk dalam *cluster* tertentu dan memungkinkan bagi setiap data yang termasuk dalam *cluster* tertentu pada suatu tahapan proses, pada tahapan berikutnya berpindah ke *cluster* yang lain [4]. Dengan kata lain, metode ini berusaha untuk meminimalkan variasi antar data yang ada di dalam suatu *cluster* dan memaksimalkan variasi dengan data yang ada di cluster lainnya.

Clustering adalah proses pengelompokan sejumlah data/objek ke dalam kelompok data sehingga setiap kelompok berisi data yang mirip [5]. Tujuan analisis kelompok adalah agar objek-objek yang bergabung dalam sebuah kelompok merupakan objek-objek yang mirip (berhubungan) satu sama lain dan berbeda (tidak berhubungan) dengan objek dalam kelompok yang lain.

Algoritma *K-Means* merupakan model *centroid*. Model *centroid* adalah model yang menggunakan *centroid* untuk membuat *cluster*. *Centroid* adalah "titik tengah" suatu *cluster*. *Centroid* berupa nilai.

Berikut ini adalah langkah-langkah algoritma *K-Means*:

1. Tentukan K sebagai jumlah *cluster* yang dibentuk.
2. Tentukan pusat (*centroid*) *cluster* awal.
Penentuan *centroid* awal dilakukan secara *random* atau acak dari objek-objek yang tersedia sebanyak k *cluster*, kemudian untuk menghitung *centroid cluster* ke-j berikutnya, digunakan rumus sebagai berikut :

$$C_i = \frac{1}{M} \sum_{j=1}^M X_j \dots\dots\dots (1)$$

keterangan :

C_i = *centroid* pada *cluster*

X_j = objek ke-j

M = banyaknya objek /jumlah objek yang menjadi anggota *cluster*

3. Perhitungan

Untuk mengukur jarak antar data dengan pusat *cluster* digunakan *Euclidian Distance*. Jarak hasil perhitungan akan dilakukan perbandingan dan dipilih jarak terdekat antar data dengan pusat *cluster*, jarak ini menggunakan bahwa data tersebut berada dalam satu kelompok dengan pusat *cluster* terdekat.

Rumusnya adalah :

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} ; i = 1, 2, 3, \dots, n \dots\dots\dots (2)$$

Dimana :

x_i : objek x ke-i

y_i : daya y ke-i

n : banyaknya objek

- 4.....Pengelompokan data

Setelah sejumlah populasi data tersebut menemukan kedekatan dengan salah satu *centroid* yang ada maka secara otomatis populasi data tersebut masuk kedalam kelas yang memiliki *centroid* yang bersangkutan.

- 5.....Langkah *iterasi*, kemudian tentukan posisi *centroid* baru dengan menggunakan persamaan.
- 6.....Ulangi langkah 3 apabila masih ada data yang berpindah kelompok, atau ada perubahan nilai *centroid* diatas nilai ambang yang ditentukan, atau juga apabila perubahan nilai pada fungsi objektif yang digunakan masih diatas nilai ambang yang ditentukan.

3. METODOLOGI PENELITIAN

3.1 Metode penelitian

Metode Penelitian merupakan urutan-urutan proses analisis data yang akan disajikan secara sistematis. Karena dengan urutan proses analisis data dapat diketahui secara cepat dan membantu pemahaman maksud dari penelitian tersebut.

3.2 Analisa Permasalahan

Museum adalah institusi permanen dalam hal melayani dan mengembangkan masyarakat, yang digunakan sebagai tempat untuk pameran tetap benda-benda warisan budaya yang bernilai luhur yang di anggap patut disimpan, seperti peninggalan sejarah, seni, dan ilmu; tempat menyimpan barang kuno. Terbuka untuk umum untuk tujuan pembelajaran, pendidikan, rekreasi, dan memberitahukan aset-aset barang berharga yang nyata dan “tidak nyata” tentang lingkungannya kepada masyarakat.

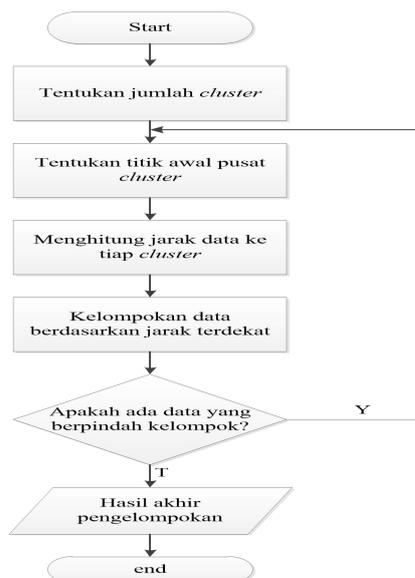
Pengunjung museum memiliki peran yang penting terhadap pengembangan museum dalam upaya mempertahankan dan meningkatkan jumlah pengunjung museum, melakukan suatu pengelompokan data pengunjung museum sangat disarankan bagi pengelola museum, untuk mengetahui Museum Deli Serdang yang memiliki jumlah pengunjung terendah dalam kunjungannya. agar pengelola museum dapat dengan mudah melakukan segala antisipasi yang akan terjadi ketika peningkatan jumlah pengunjung maupun ketika penurunan jumlah pengunjung. Jika terjadi peningkatan jumlah pengunjung maka pengelola museum dapat mempersiapkan saran dan prasarana yang baik bagi pengunjung sehingga pengunjung merasa nyaman ketika berada di museum. Namun jika terjadi penurunan jumlah pengunjung maka pengelola museum dapat melakukan suatu upaya untuk mempromosikan museum secara menarik agar pengunjung tertarik. Untuk itu, Pengelompokan tersebut dapat menggunakan metode pengelompokan Algoritma *K-Means*.

3.3 Algoritma Sistem

Algoritma Sistem merupakan penjelasan langkah-langkah penyelesaian masalah dalam perancangan Data Mining dalam Pengelompokan pengunjung di Museum daerah Deli Serdang. Berikut ini algoritma sistem pada penyelesaian sistem pengelompokan pengunjung menggunakan metode *K-Means*.

3.3.1 Flowchart Algoritma *K-Means Clustering*

Flowchart digunakan untuk memperinci rencana pembangunan yang akan dijalankan pada pengelompokan pengunjung Museum daerah Deli Serdang yang akan menggunakan metode Algoritma *K-Means Clustering* yaitu sebagai berikut :



Gambar 3.1 *Flowchart Algoritma K-Means Clustering*

3.3.2 Pengolahan Data menggunakan Algoritma *K-Means*

Berikut ini adalah data yang digunakan sebagai sampel dalam penelitian ini yaitu sebagai berikut:

Tabel 3.1 Data Pengunjung Museum

No	Bulan	Pengunjung				
		TK/PAUD	SD	SMP/MTS	SMK/SMK/ MAN	UNIVERSITAS
1	Sep-18	4	0	34	23	59
2	Oct-18	83	2051	381	75	6
3	Nov-18	144	11309	991	227	38
4	Dec-18	349	5925	831	115	4
5	Jan-19	202	3027	301	76	10
6	Feb-19	194	7092	471	134	12
7	Mar-19	205	7102	450	155	30
8	Apr-19	220	7110	443	148	8
9	May-19	202	3062	320	80	15
10	June 2019	198	1860	479	130	12
11	Jul-19	760	63	70	71	0
12	Aug-19	0	1810	4	300	6
13	Sep-19	107	1452	0	2	9
14	Oct-19	242	7934	496	65	3
15	Nov-19	146	6815	482	421	37
16	Dec-19	295	2484	780	26	30

Setelah mendapatkan data maka melakukan perhitungan sebagai berikut :

1. Menentukan Titik Pusat *Cluster*

Menentukan banyaknya *Cluster* (*K*) yang akan dibuat yaitu 3 *cluster*. Setelah menentukan banyaknya *cluster*, tentukan titik pusat awal *cluster* (*centroid*) secara acak (random). Berikut ini adalah titik pusat (*centroid*) yang telah dipilih.

Tabel 3.2 Titik Pusat Awal *Cluster*

Bulan	Pengunjung				
	TK/PAUD	SD	SMP/MTS	SMA/SMK/ MAN	UNIVERSITAS
Oct-18	83	2051	381	75	6
Jan-19	202	3027	301	76	10
Jun-19	198	1860	479	130	12

2. Perhitungan jarak dengan setiap pusat *cluster* dengan rumus :

$$D(x_2, x_1) = \|x_2 - x_1\|_2 = \sqrt{\sum_{j=1}^p |x_2 - x_{1j}|^2}$$

3. Menghitung jarak antara variabel dari setiap sampel data dengan setiap titik pusat.

4. Iterasi ke-1

Tabel 3.3 Hasil Perhitungan Iterasi 1

No	Bulan	Pusat m1	Pusat m2	Pusat m3	Jarak terdekat	Cluster
1	Sep-18	2082.97	3046.05	1925.86	1925.86	3
2	Oct-18	0.00	986.49	249.74	0.00	1
3	Nov-18	9279.58	8312.31	9463.55	8312.31	2
4	Dec-18	3909.31	2949.99	4083.04	2949.99	2
5	Jan-19	986.49	0.00	1181.74	0.00	2
6	Feb-19	5043.37	4068.97	5232.01	4068.97	2
7	Mar-19	5053.63	4078.54	5242.18	4078.54	2
8	Apr-19	5061.76	4086.14	5250.20	4086.14	2

9	May-19	1019.86	40.34	1213.51	40.34	2
10	Jun2019	249.74	1181.74	0.00	0.00	3
11	Jul-19	2123.03	3024.92	1927.68	1927.68	3
12	Aug-19	507.67	1288.52	544.30	507.67	1
13	Sep-19	714.06	1608.02	648.52	648.52	3
14	Oct-19	5886.28	4911.05	6074.54	4911.05	2
15	Nov-19	4778.13	3808.49	4963.87	3808.49	2
16	Dec-19	628.18	732.01	707.48	628.18	1

a. Hitung nilai WCV (*Within Cluster Variation*) dengan cara memangkatkan jarak terdekat *cluster* dan menjumlahkan setiap nilai WCV.

$$WCV = 2082.97^2 + 0^2 + 8312.31^2 \dots + 628.18^2$$

$$WCV = 174807099$$

b. Hitung nilai BCV (*Between Cluster Variation*) dengan cara menjumlahkan hasil dari jarak diantara setiap *centroid*.

$$d(m1,m2) = \sqrt{(m1 - m2)^2} = 986.49$$

$$d(m1,m3) = \sqrt{(m1 - m3)^2} = 249.74$$

$$d(m2,m3) = \sqrt{(m2 - m3)^2} = 1181.74$$

$$BCV = d(m1,m2) + d(m1,m3) + d(m2,m3) = 2417.97$$

c. Menghitung nilai besar rasio dengan membandingkan nilai BCV dan WCV

$$BCV/WCV = 2417.97/174807099$$

$$= 0.000013$$

Kemudian langkah berikutnya pada iterasi ke-2 adalah :

5. Menghitung kembali pusat-pusat baru dari kelompok (*cluster*) yang baru terbentuk

Tabel 3.4 Titik *cluster* Baru

No	Centroid	TK/PAUD	SD	SMP/MTS	SMA/SMK/ MAN	UNIVERSTITAS
1	m1	126	2115	388.33	133.67	14
2	m2	211.56	6597.33	531.67	157.89	17.44
3	m3	267.25	843.75	145.75	56.5	20

6. Menghitung kembali jarak antara variabel dari setiap sampel data dengan *centroid*.

Tabel 3.5 Hasil Perhitungan Iterasi 2

No	Bulan	Pusat m1	Pusat m2	Pusat m3	Jarak terdekat	Cluster
1	Sep-18	2151.26	6620.83	892.38	892.38	3
2	Oct-18	97.49	4551.41	1243.90	97.49	1
3	Nov-18	9214.25	4735.04	10501.45	4735.04	2
4	Dec-18	3842.17	750.02	5128.26	750.02	2
5	Jan-19	921.13	3578.73	2189.84	921.13	1
6	Feb-19	4978.15	499.29	6257.62	499.29	2
7	Mar-19	4988.08	511.44	6266.73	511.44	2
8	Apr-19	4996.21	520.53	6274.15	520.53	2
9	May-19	954.01	3542.53	2226.17	954.01	1
10	Jun-19	280.08	4737.73	1074.28	280.08	1
11	Jul-19	2172.12	6574.13	926.67	926.67	3
12	Aug-19	533.24	4823.07	1041.46	533.24	1
13	Sep-19	779.80	5176.14	648.06	648.06	3
14	Oct-19	5821.57	1340.79	7098.97	1340.79	2
15	Nov-19	4709.80	351.79	5993.06	351.79	2
16	Dec-19	574.44	4123.79	1759.12	574.44	1

- a. Hitung nilai WCV (*Within Cluster Variation*) dengan cara memangkatkan jarak terdekat *cluster* dan menjumlahkan setiap nilai WCV.

$$\begin{aligned} \text{WCV} &= 892.38^2 + 97.49^2 + 4735.04^2 \dots + 574.44^2 \\ &= 30222347.54 \end{aligned}$$

- b. Hitung nilai BCV (*Between Cluster Variation*) dengan cara menjumlahkan hasil dari jarak diantara setiap *centroid*.

$$\begin{aligned} d(m1,m2) &= \sqrt{(m1 - m2)^2} = 4485.50 \\ d(m1,m3) &= \sqrt{(m1 - m3)^2} = 1304.17 \\ d(m2,m3) &= \sqrt{(m2 - m3)^2} = 5767.67 \\ \text{BCV} &= d(m1,m2) + d(m1,m3) + d(m2,m3) = 11557.35 \end{aligned}$$

- c. Menghitung nilai besar rasio dengan membandingkan nilai BCV dan WCV

$$\begin{aligned} \text{BCV/WCV} &= 11557.35/30222347.54 \\ &= 0.000038 \end{aligned}$$

Bila dibandingkan maka rasio sekarang (0.00038) lebih besar dari rasio sebelumnya (0.00001) oleh karena itu algoritma dilanjutkan kelangkah berikutnya.

Kemudian langkah berikutnya pada iterasi ke-3 adalah :

7. Menghitung kembali pusat-pusat baru dari kelompok (*cluster*) yang baru terbentuk.

Tabel 3.6 Titik *Cluster* Baru

No	<i>Centroid</i>	TK/PAUD	SD	SMP/MTS	SMA/SMKMAN	UNIVERSITAS
1	m1	163.33	2382.33	377.5	114.5	13.17
2	m2	214.29	7612.42	594.86	180.71	18.86
3	m3	290.33	505	34.67	32	22.67

8. Menghitung kembali jarak antara variabel dari setiap sampel data dengan *centroid*.

Tabel 3.7 Hasil Perhitungan Iterasi 3

No	Bulan	Pusat m1	Pusat m2	Pusat m3	Jarak terdekat	<i>Cluster</i>
1	Sep-18	2414.40	7637.68	581.73	581.73120	3
2	Oct-18	343.30	5568.10	1598.49	343.30218	1
3	Nov-18	8948.49	3718.75	10848.99	3718.75	2
4	Dec-18	3576.41	1710.51	5479.16	1710.51	2
5	Jan-19	651.49	4596.04	2537.97	651.49	1
6	Feb-19	4710.74	537.42	6602.93	537.42	2
7	Mar-19	4720.61	531.40	6611.76	531.40	2
8	Apr-19	4728.58	526.03	6619.02	526.03	2
9	May-19	684.07	4559.84	2574.85	684.07	1
10	Jun-19	533.46	5753.84	1432.37	533.45525	1
11	Jul-19	2414.94	7588.11	647.49	647.48517	3
12	Aug-19	726.78	5837.59	1363.95	726.77522	1
13	Sep-19	1011.86	6192.59	965.77	965.76797	3
14	Oct-19	5553.72	357.20	7443.57	357.20	2
15	Nov-19	4444.58	843.41	6339.44	843.41	2
16	Dec-19	444.74	5134.74	2114.73	444.74	1

- a. Hitung nilai WCV (*Within Cluster Variation*) dengan cara memangkatkan jarak terdekat *cluster* dan menjumlahkan setiap nilai WCV.

$$\begin{aligned} \text{WCV} &= 581.73^2 + 343.30^2 + 3718.75^2 \dots + 444.74^2 \\ \text{WCV} &= 22152938.76 \end{aligned}$$

- b. Hitung nilai BCV (*Between Cluster Variation*) dengan cara menjumlahkan hasil dari jarak diantara setiap *centroid*.

$$\begin{aligned} d(m1,m2) &= \sqrt{(m1 - m2)^2} = 5235.27 \\ d(m1,m3) &= \sqrt{(m1 - m3)^2} = 1914.40 \end{aligned}$$

$$d(m2,m3) = \sqrt{(m2 - m3)^2} = 7131.42$$

$$BCV = d(m1,m2) + d(m1,m3) + d(m2,m3) = 14281.09$$

c. Menghitung nilai besar rasio dengan membandingkan nilai BCV dan WCV

$$BCV/WCV = 14281.09/22152938.76 = 0.00064$$

Bila dibandingkan maka rasio sekarang (0.00064) lebih besar dari rasio sebelumnya (0.00038) oleh karena itu algoritma dilanjutkan kelangkah berikutnya.

Kemudian langkah berikutnya pada iterasi ke-4 adalah :

9. Menghitung kembali pusat-pusat baru dari kelompok (*cluster*) yang baru terbentuk :

Tabel 3.8 Titik *Cluster* Baru

No	Centroid	TK/PAUD	SD	SMP/MTS	SMA/SMKMAN	UNIVERSITAS
1	m1	163.33	2382.33	377.5	114.5	13.17
2	m2	214.29	7612.42	594.86	180.71	18.86
3	m3	290.33	505	34.67	32	22.67

Tabel 3.9 Hasil Perhitungan Iterasi 4

No	Bulan	Pusat m1	Pusat m2	Pusat m3	Jarak terdekat	Cluster
1	Sep-18	2414.40	7637.68	581.73	581.73120	3
2	Oct-18	343.30	5568.10	1598.49	343.30218	1

Tabel 3.9 Hasil Perhitungan Iterasi 4

No	Bulan	Pusat m1	Pusat m2	Pusat m3	Jarak terdekat	Cluster
1	Sep-18	2414.40	7637.68	581.73	581.73120	3
2	Oct-18	343.30	5568.10	1598.49	343.30218	1
3	Nov-18	8948.49	3718.75	10848.99	3718.75	2
4	Dec-18	3576.41	1710.51	5479.16	1710.51	2
5	Jan-19	651.49	4596.04	2537.97	651.49	1
6	Feb-19	4710.74	537.42	6602.93	537.42	2
7	Mar-19	4720.61	531.40	6611.76	531.40	2
8	Apr-19	4728.58	526.03	6619.02	526.03	2
9	May-19	684.07	4559.84	2574.85	684.07	1
10	Jun-19	533.46	5753.84	1432.37	533.45525	1
11	Aug-19	726.78	5837.59	1363.95	726.77522	1
12	Sep-19	1011.86	6192.59	965.77	965.76797	3
13	Oct-19	5553.72	357.20	7443.57	357.20	2
14	Nov-19	4444.58	843.41	6339.44	843.41	2
15	Dec-19	444.74	5134.74	2114.73	444.74	1
16	Jun-19	533.46	5753.84	1432.37	533.45525	1

10. Menghitung kembali jarak antara variabel dari setiap sampel data dengan *centroid*.

a. Hitung nilai WCV (*Within Cluster Variation*) dengan cara memangkatkan jarak terdekat *cluster* dan menjumlahkan setiap nilai WCV.

$$WCV = 581.73^2 + 343.30^2 + 3718.75^2 \dots + 444.74^2 = 22152938.76$$

b. Hitung nilai BCV (*Between Cluster Variation*) dengan cara menjumlahkan hasil dari jarak diantara setiap *centroid*.

$$d(m1,m2) = \sqrt{(m1 - m2)^2} = 5235.27$$

$$d(m1,m3) = \sqrt{(m1 - m3)^2} = 1914.40$$

$$d(m2,m3) = \sqrt{(m2 - m3)^2} = 7131.42$$

$$BCV = d(m1,m2) + d(m1,m3) + d(m2,m3) = 14281.09$$

c. Menghitung nilai besar rasio dengan membandingkan nilai BCV dan WCV

$$BCV/WCV = 14281.09/22152938.76$$

$$= 0.00064$$

Bila dibandingkan maka rasio sekarang (0.00064) sudah tidak lagi lebih besar dari rasio sebelumnya (0.00064) oleh karena itu algoritma dihentikan. Setelah hasil perhitungan selesai, maka akan dilihat bahwa pengelompokan data pada setiap *cluster* tidak akan berubah. Kelompok dari setiap *cluster* pada iterasi ke-4 yaitu dapat dilihat pada tabel berikut ini :

Tabel 3.10 Tabel kelompok *cluster* 1 pada iterasi 4

No	Bulan	TK/PAUD	SD	SMP/MTS	SMA/SMK/MAN	UNIVERSITAS
1	Oct-18	83	2051	381	75	6
2	Jan-19	202	3027	301	76	10
3	May-19	202	3062	320	80	15
4	Jun-19	198	1860	479	130	12
5	Aug-19	0	1810	4	300	6
6	Dec-19	295	2484	780	26	30

Selanjutnya adalah tabel kelompok *cluster* 2 pada iterasi 4, seperti terlihat pada tabel dibawah ini :

Tabel 3.11 Tabel kelompok *cluster* 2 pada iterasi 4

No	Bulan	TK/PAUD	SD	SMP/MTS	SMA/SMK/MAN	UNIVERSTITAS
1	Nov-18	144	11309	991	227	38
2	Dec-18	349	5925	831	115	4
3	Feb-19	194	7092	471	134	12
4	Mar-19	205	7102	450	155	30
5	Apr-19	220	7110	443	148	8
6	Oct-19	242	7934	496	65	3
7	Nov-19	146	6815	482	421	37

Selanjutnya adalah tabel kelompok *cluster* 3 pada iterasi 4, seperti terlihat pada tabel dibawah ini :

Tabel 3.12 Tabel kelompok *cluster* 3 pada iterasi 4

No	Bulan	TK/PAUD	SD	SMP/MTS	SMA/SMK/MAN	UNIVERSITAS
1	Sep-18	4	0	34	23	59
2	Jul-19	760	63	70	71	0
3	Sep-19	107	1452	0	2	9

Untuk melakukan penilaian terhadap pengelompokan jumlah kunjungan pada Museum daerah Deli Serdang dapat diterapkan dengan metode *K-Means Clustering*. Data diolah dengan melakukan *K-Means* yang dikelompokkan menjadi 3 *cluster* yaitu *cluster* tinggi (C1) yakni jumlah kunjungan pengunjung tinggi, C2= jumlah kunjungan pengunjung sedang, C3= jumlah kunjungan pengunjung rendah. Diperoleh hasil pengelompokan C1= Bulan Oktober 2018, Januari 2019, Mei 2019, Agustus 2019, Desember 2019, C2= Bulan November 2018, Desember 2018, Februari 2019, Maret 2019, April 2019, Oktober 2019, November 2019. C3= Bulan September 2018, Juli 2019, September 2019.

3.4 Hasil

Adapun hasil pengujian sistem di dalam Penerapan Data Mining Untuk Pengelompokan data pengunjung museum adalah sebagai berikut :

1. Input data pengunjung museum , berikut inputannya :



Gambar 5.1 Penginputan Data Pengunjung Museum

2. Input proses *clustering* pada data pengunjung museum, berikut proses penginputannya:



Gambar 5.2 Proses *Clustering* Data Pengunjung Museum

3. Hasil *cluster* data pengunjung berdasarkan dari pengelompokan dengan menggunakan metode *K-Means Clustering*, berikut adalah hasil :



Gambar 5.3 Hasil *cluster* data pengunjung museum

4. Laporan berdasarkan dari pengelompokan dengan menggunakan metode *K-Means Clustering*, berikut adalah hasil laporannya:

MUSEUM DELI SERDANG						
Kompleks Perkantoran Pemerintahan Desa Tanjung Gabus I Lubuk Pakam, Deli Serdang Sumatera Utara						
Laporan Hasil Analisa Pengelompokan Data Pengunjung Museum						
Kode Pengunjung	Bulan	TK/PAUD	SD	SMP/MTS	SMA/SMK/MA	cluster
9	Mei 2019	202	3062	320	80	15 1
2	Oktober 2018	83	2051	381	75	0 1
12	Agustus 2018	0	1810	4	300	6 1
5	Januari 2019	202	3027	301	76	10 1
10	Juni 2019	198	1860	479	134	12 1
15	Desember 2019	295	2484	780	26	30 1
3	November 2018	144	11309	991	227	38 2
4	Desember 2018	349	8928	831	118	4 2
6	Februari 2019	194	7092	471	134	12 2
8	Agustus 2019	220	7110	443	148	8 2
14	Oktober 2019	242	7934	490	69	3 2
18	November 2019	148	8515	452	421	37 2
7	Maret 2019	208	7102	480	186	30 2
1	September 2018	4	0	34	23	89 3
11	Juli 2019	760	63	70	71	0 3
13	September 2019	107	1452	0	2	9 3

Lubuk Pakam, 6/24/2020
Diketahui Oleh,

5.9 Laporan data pengunjung museum

4. KESIMPULAN

Berdasarkan analisa pada permasalahan yang terjadi dalam kasus yang diangkat tentang pengelompokan data pengunjung Museum Deli Serdang menggunakan algoritma *K-Means Clustering*, maka dapat ditarik kesimpulan sebagai berikut :

1. Data mining sangat berpengaruh dalam pengelompokan data pengunjung museum sehingga membantu pengelola museum dalam meningkatkan jumlah pengunjung Museum Deli Serdang.
2. Algoritma *K-Means Clustering* melakukan penilaian terhadap pengelompokan jumlah kunjungan pada Museum daerah Deli Serdang dapat diterapkan dengan metode *K-Means Clustering*. Data diolah dengan melakukan *K-Means* yang dikelompokkan menjadi 3 *cluster* yaitu *cluster* tinggi (C1) yakni jumlah kunjungan pengunjung tinggi, *cluster* sedang (C2) yakni jumlah kunjungan pengunjung sedang dan *cluster* rendah (C3) yakni jumlah kunjungan pengunjung rendah.
3. Penerapan sistem yang dirancang dengan menggunakan aplikasi *Data Mining* menggunakan Metode Algoritma *K-Means Clustering* dengan cara mempersiapkan perangkat lunak dan perangkat keras lalu mengoperasikan sistem yang telah dibangun.

UCAPAN TERIMA KASIH

Alhamdulillah, puji syukur kehadirat Allah SWT atas izin-Nya yang telah melimpahkan rahmat, taufik dan hidayah-Nya sehingga dapat menyelesaikan jurnal ilmiah ini dengan baik. Ucapan terima kasih istimewa ditujukan kepada orang tua, yang telah mengasuh, membesarkan dan selalu memberikan doa, motivasi serta pengorbanan baik bersifat moril maupun materil yang tidak terhingga selama menjalani pendidikan. Ucapan terima kasih yang sebesar-besarnya juga ditujukan terutama kepada Bapak Rudi Gunawan, SE., M.Si., selaku Ketua Sekolah Tinggi Manajemen Informatika Dan Komputer (STMIK) Triguna Dharma Medan. Bapak Dr.Zulfian Azmi, ST., M.Kom., selaku Wakil Ketua I Bidang Akademik STMIK Triguna Dharma Medan. Bapak Marsono, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi STMIK Triguna Dharma Medan. Ibu Yohanni Syahra, S.Si., M.Kom., selaku Dosen Pembimbing I yang telah memberikan saran, arahan dan dukungannya serta motivasi, sehingga penelitian ini dapat terselesaikan dengan baik dan tepat waktu. Ibu Ita Mariami, S.E., M.Si., selaku Dosen Pembimbing II yang telah memberikan bimbingan tata cara penulisan, saran dan motivasi sehingga penelitian ini dapat terselesaikan dengan baik dan tepat waktu. Seluruh Dosen, Staff dan Pegawai di STMIK Triguna Dharma Medan. Bapak Afri Deliansyah Nasution, S.Pd, selaku Kasi Cagar Budaya dan Museum yang telah mengizinkan melakukan Riset guna memenuhi data dan bahan yang dibutuhkan dalam menyelesaikan kasus yang diangkat dan seluruh teman-teman di STMIK Triguna Dharma Medan yang telah berbagi dalam suka maupun duka dan membantu hingga terselesaikannya penelitian ini.

REFERENSI

- [1] J. P. Sejarah, R. Sosial, And H. Volume, "Peran Museum Dalam Pembelajaran Sejarah," Vol. 2, Pp. 10–20, 2019.
- [2] F. L. Sibuea, A. Sapta, S. Informasi, And S. Royal, "Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering," Vol. Iv, No. 1, 2017.
- [3] L. Bruno, "Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means," *J. Chem. Inf. Model.*, Vol. 53, No. 9, Pp. 1689–1699, 2019, Doi: 10.1017/Cbo9781107415324.004.
- [4] D. Triyansyah And D. Fitrihanah, "Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing," No. September, 2018, Doi: 10.22441/Incomtech.V8i2.4174.

[5] Vulandari, Retno Tri. 2017. *Data Mining Teori Dan Aplikasi Rapidminer*. Surakarta : Gava Media.

BIOGRAFI PENULIS

	<table border="1"> <tr> <td data-bbox="639 450 815 501">Nama :</td> <td data-bbox="815 450 1423 501">NATASAH LIAZA</td> </tr> <tr> <td data-bbox="639 501 815 553">Email :</td> <td data-bbox="815 501 1423 553">natasahliaza25@gmail.com</td> </tr> <tr> <td data-bbox="639 553 815 604">T.T.L</td> <td data-bbox="815 553 1423 604">Tanjung Morawa, 25 Januari 1999</td> </tr> <tr> <td data-bbox="639 604 815 725">Program Studi :</td> <td data-bbox="815 604 1423 725">Sistem Informasi</td> </tr> <tr> <td data-bbox="639 725 815 831">Mobile :</td> <td data-bbox="815 725 1423 831">0896-8946-3998</td> </tr> </table>	Nama :	NATASAH LIAZA	Email :	natasahliaza25@gmail.com	T.T.L	Tanjung Morawa, 25 Januari 1999	Program Studi :	Sistem Informasi	Mobile :	0896-8946-3998
Nama :	NATASAH LIAZA										
Email :	natasahliaza25@gmail.com										
T.T.L	Tanjung Morawa, 25 Januari 1999										
Program Studi :	Sistem Informasi										
Mobile :	0896-8946-3998										
	<table border="1"> <tr> <td data-bbox="639 891 815 943">Nama :</td> <td data-bbox="815 891 1423 943">YOHANNI SYAHRA, S.Si., M.Kom</td> </tr> </table>	Nama :	YOHANNI SYAHRA, S.Si., M.Kom								
Nama :	YOHANNI SYAHRA, S.Si., M.Kom										
	<table border="1"> <tr> <td data-bbox="639 1323 815 1375">Nama :</td> <td data-bbox="815 1323 1423 1375">ITA MARIAMI, S.E., M.Si</td> </tr> </table>	Nama :	ITA MARIAMI, S.E., M.Si								
Nama :	ITA MARIAMI, S.E., M.Si										