

Analisis Data Pengunjung Mall Nipah Mendukung Strategi Digital Marketing Menggunakan *K-Means Clustering*

Charos George Selan^{*1}, Nikmah Jayanti², Medy Wisnu Prihatmono³

^{1,2,3}Sistem Informasi, STMIK Profesional Makassar, Indonesia
Email: ¹charos.george09@gmail.com, ²nikmahjayanti44@gmail.com,
³medywisnuprihatmonoclay@gmail.com

Abstrak

Pada era saat ini system digital marketing berkembang dengan sangat pesat, perusahaan berlomba dalam menyusun strategi digital untuk menarik para pengunjung. nipah mall merupakan mall berkonsepkan "green bulding" ruang terbuka hijau dengan fasilitas terbaik dan kenyamanan para pengunjung. Pada strategi kali ini system digital marketing membuat suatu event yang dimana para pengunjung mengikuti petunjuk yang telah di sebarakan sebagaimana pada event ini mengacu pada perlombaan menjawab sebuah pertanyaan (Quis) dalam aplikasi berbasis barcode untuk mendapatkan sebuah voucher belanja. Penelitian ini bertujuan untuk menganalisa data pengunjung yang dimana penerapan strategi telah dilakukan dan sebagai upaya peningkatan volume pengunjung untuk mengetahui: [1] jumlah respondent pengunjung [2] jumlah respondent yang mengikuti quis dengan menjawab benar dan salah. Berdasarkan uraian tersebut penelitian dilakukan untuk menghasilkan kesimpulan penerapan strategi marketing digital ini terlaksana dengan baik oleh devisi digital marketing & communication. Pada pengelompokan data tersebut dapat dilihat karakteristiknya sehingga diketahui cluster rendah, sedang dan tinggi, untuk menganalisa data tersebut selama pelaksanaan event ini perlu di analisa menggunakan metode *K-Means Clustering*. Algoritma *K-Means* berperan untuk menyelesaikan penelitian ini dengan metode pengelompokan data Sehingga menghasilkan simulasi pengetahuan data pengunjung mengikuti event tersebut, dan mengimplementasikan hasil *K-Means* dalam Rapid Miner agar dapat melihat hasil visualisasi algoritma *K-Means*.

Kata kunci: Algoritma *K-Means*, *Clustering*, *Rapid Miner*, *Stategi Marketing*

Abstract

In the current era, the digital marketing system is developing very rapidly, companies are competing in developing digital strategies to attract visitors. nipah mall is a mall with the concept of "green bulding" green open space with the best facilities and the convenience of visitors. In this strategy, the digital marketing system creates an event where visitors follow the instructions that have been distributed as in this event it refers to a competition to answer a question (Quis) in a barcode-based application to get a shopping voucher. This study aims to analyze visitor data where the implementation of the strategy has been carried out and as an effort to increase the volume of visitors to find out: [1] the number of visitor respondents [2] the number of respondents who took the quiz with correct and incorrect answers. Based on this description, research was conducted to conclude that the implementation of this digital marketing strategy was carried out well by the digital marketing & communication division. The characteristics of the data grouping can be seen so that it is known that the clusters are low, medium and high, to analyze the data during the implementation of this event, it is necessary to analyze it using the *K-Means Clustering* method. The *K-Means* algorithm plays a role in completing this research with the data grouping method so as to produce a simulation of visitor data knowledge following the event, and implement the *K-Means* results in *Rapid Miner* in order to see the results of the *K-Means* algorithm visualization..

Keywords: *Clustering*, *K-Means Algorithm*, *Marketing Strategy*, *Rapid Miner*

1. PENDAHULUAN

Pada era saat ini system digital marketing berkembang dengan sangat pesat karena banyaknya persaingan marketing disetiap perusahaan dan berlomba dalam Menyusun strategi digital untuk menarik para pengunjung. Nipah mall merupakan mall yang berkonsep "green bulding" dengan menawarkan

konsep ruang terbuka hijau dengan fasilitas terbaik dan kenyamanan para pengunjung yang datang. Pada strategi kali ini system digital marketing membuat suatu event yang dimana para pengunjung mengikuti petunjuk-petunjuk yang telah di sebarakan sebagaimana pada event ini mengacu pada perlombaan menjawab sebuah pertanyaan (quis) dalam aplikasi berbasis barcode untuk mendapatkan sebuah voucher belanja.

Penelitian ini bertujuan untuk menganalisa data pengunjung yang dimana penerapan strategi telah dilakukan dan sebagai upaya peningkatan volume pengunjung untuk mengetahui : 1) jumlah respondent pengunjung, 2) jumlah respondent yang mengikuti quis dengan menjawab benar dan salah.

Untuk menyelesaikan permasalahan tersebut dan mempermudah proses pengolahan data yang cukup banyak, yaitu perlu adanya suatu system pengambilan pengelompokan data untuk mengetahui keputusan dalam analisis untuk mengetahui jumlah respondent pengunjung dan jumlah respondent yang mengikuti event tersebut yang menjawab benar dan salah, Teknik yang digunakan adalah data mining untuk mengolah data yang berskala besar.

Data mining adalah suatu Teknik pengolahan data yang berskala besar atau banyak, dan kemudian data yang telah diolah dengan Teknik data mining akan menghasilkan keputusan, dan pada data mining mempunyai beberapa metode yaitu *classification*, *clustering*, *estimation*, *prediction*, *association*. salah satu metode yang digunakan pada penelitian ini adalah *clustering* (Sugianto et al., 2020).

Clustering adalah metode yang ada pada data mining dengan Teknik pengumpulan data berdasarkan persamaan karakteristik antara data satu dengan data yang lainnya, algoritma yang digunakan pada *clustering* data mining adalah Algoritma *K-Means*, *K-Means* merupakan algoritma pengelompokan data berdasarkan hasil penetapan yang non hirarki yang berusaha untuk mempartisi objek yang ada kedalam satu atau lebih *cluster* atau kelompok objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu *cluster* yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan kedalam *cluster* yang lain (Sugianto et al., 2020).

Berdasarkan uraian tersebut penelitian dilakukan untuk menghasilkan kesimpulan penerapan strategi marketing digital ini terlaksana dengan baik dan memudahkan devisi digital marketing & communication mengetahui volume pengunjung, jumlah respondent pengunjung mall dan pengelompokan data yang menjawab benar dan salah.

pengelompokan tersebut dapat dilihat karakteristiknya sehingga diketahui *cluster* rendah, *cluster* sedang dan *cluster* tinggi untuk dapat menganalisa data tersebut selama pelaksanaan event ini dan perlu adanya Analisa lebih dalam menggunakan metode *K-Means Clustering*.

Algoritma *K-Means* berperan untuk menyelesaikan penelitian ini dengan metode pengelompokan data Sehingga menghasilkan simulasi pengetahuan data pengunjung mengikuti event tersebut, dan mengimplementasikan hasil *K-Means* dalam Rapidminer agar dapat melihat hasil visualisasi algoritma *k-means*.

1.1. Analisis Data

Analisis data merupakan kegiatan yang meliputi proses berpikir untuk merinci dan menguraikan suatu keseluruhan yang dijadikan komponen sehingga setiap komponen lebih mudah untuk dimengerti, baik itu hubungan antar komponen, serta fungsi dari masing-masing komponen ataupun fungsi secara keseluruhan, Komaruddin (2001). Analisis data adalah proses menganalisis data termasuk menafsirkan data yang sudah dilakukan analisa. Selain itu, analisis data juga termasuk teknik pengumpulan data karena proses pengumpulan data juga menentukan seberapa akurat data yang akan digunakan. Jhoney Turkey (2011) dalam (Triyansyah & Fitriana, 2018).

1.2. Clustering

Clustering adalah salah satu teknik data mining yang membentuk kelompok dari objek-objek yang mempunyai karakteristik pada data yang besar. Tujuan Teknik clustering adalah pengelompokan sejumlah data atau objek ke dalam cluster sehingga pada setiap cluster akan berisi data yang sama. clustering melakukan pengelompokan data yang didasarkan pada kesamaan antara objek, oleh karena itu klasterisasi digolongkan sebagai metode unsupervised learning (Santosa, 2007) dalam (Anggreini, 2019).

1.3. Clustering

Strategi adalah sebuah rencana (Plan) yang disatukan, menyeluruh, dan terpadu yang menghubungkan strategi perusahaan dengan tantangan lingkungan dan yang dirancang untuk memastikan bahwa tujuan utama perusahaan dapat dicapai melalui pelaksanaan yang tepat dalam perusahaan, Willy Pratama Widharta dan Sugiono Sugiharto(2013)[3]. Strategi merupakan suatu proses penentuan rencana para pemimpin puncak yang berfokus pada tujuan jangka panjang organisasi, disertai penyusunan suatu cara atau upaya pengunjung mengikuti event tersebut, dan mengimplementasikan hasil *K-Means* dalam Rapidminer agar dapat melihat hasil visualisasi algoritma k-means (Willy Pratama Widharta dan Sugiono Sugiharto, S.E., 2012).

1.4. Digital Marketing

Digital Marketing ialah kegiatan marketing termasuk branding yang menggunakan berbagai media berbasis web seperti blog, website, e-mail, adwords, ataupun jejaring sosial. Tentu saja digital marketing bukan hanya berbicara tentang marketing internet (Ridwan Sanjaya dan Josua Tarigan (2009:47) dalam (Dimas Hendika Wibowo, 2015).

2. METODE PENELITIAN

2.1. Algoritma K-means

K-Means salah satu algoritma pengelompokan data (*cluster*) non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster*/kelompok. tujuan dari data *clustering* ini adalah untuk meminimalisasikan objective function yang diset dalam proses *clustering*, yang pada umumnya berusaha meminimalisasikan variasi di dalam suatu *cluster* dan memaksimalkan variasi antar *cluster*. Adapun Langkah-Langkah pada pengelompokan algoritma *K-Means* (Purnamasari, 2011):

- Langkah 1 : Tentukan secara acak Nilai k sebagai jumlah *cluster* yang ingin dibentuk. tetapkan pusat *cluster*
- Langkah 2 : Menghitung Jarak antara data dan pusat *cluster* dihitung menggunakan Euclidian Distance. Untuk menghitung jarak semua data ke setiap titik pusat *cluster* dapat menggunakan teori jarak Euclidean. Dituliskan pada *equation* (1) :

$$D_{i,j} = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + (X_{3i} - X_{3j})^2 + \dots + (X_{Ki} - X_{Kj})^2} \quad (1)$$

Keterangan:

$D(i,j)$: Jarak data ke i ke pusat *cluster* j

X_{ki} : Data ke i pada atribut data ke k

X_{kj} : Titik pusat ke j pada atribut ke k

- Langkah 3 : kelompokkan data ke dalam *cluster* yang dengan jarak yang paling pendek menggunakan persamaan. Dituliskan pada *equation* (2) :

$$\text{Min} \sum_k^k = 1 \text{ dik} = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + (X_{3i} - X_{3j})^2 + \dots + (X_{Ki} - X_{Kj})^2} \quad (2)$$

- Langkah 4 : Hitung pusat *cluster* baru menggunakan persamaan. Dituliskan pada *equation* (3) :

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p} \quad (3)$$

Keterangan :

X_{ij} : E kluster ke-k

P : banyaknya anggota kluster ke k

- Langkah 4 : Proses penentuan pusat *cluster* dan penempatan data dalam *cluster* diulangi sampai

nilai centroid tidak berubah lagi (Purnamasari, 2011).

2.2. Pengumpulan Data

Pada tahap pengumpulan data ini menerapkan algoritma K-Means dalam menentukan analisis data pengunjung mall nipah dengan mendapatkan data dari aplikasi dashboard data admin (data member), <https://nipah.id/kuisramadan>, yang dimana jumlah data respondent pengunjung mall yang mengikuti event quis ramdhan ini Data yang digunakan dalam penelitian ini yaitu jumlah Reshpondent pengunjung mall mengikuti event quis Ramadhan selama priode 1-30 hari/ puasa Ramadhan yang terdiri dari 964 reshpondent. Data yang akan diolah menggunakan *clustering* meliputi 3 *cluster* yaitu jumlah reshpondent, jumlah reshponden jawab benar dan salah.

2.3. Pengolahan Data

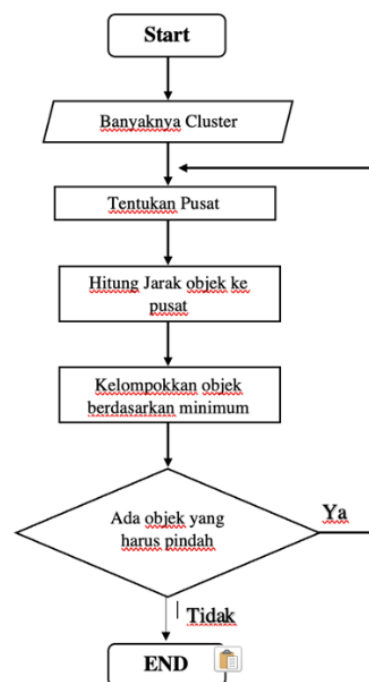
Pada pengolahan data ini data yang telah diperoleh pada dashboard admin (data member) akan diolah terlebih dahulu dan setelah diolah akan dilakukan *clustering*, dan data reshpondent akan dijumlah akan dijumlahkan setiap kriterianya yaitu pada tahapan pengolahan data akan didapatkan kalkulasi nilai yang nantinya akan dioperasikan dalam selanjutnya atau tahap *clustering* (Sianipar et al., 2020). Data Reshpondent Pengunjung dapat dilihat pada Tabel 1.

Tabel 1. Data Responden Pengunjung

ID_PERTANYAAN	Jumlah Respon	Jumlah respondent	
		benar	salah
Ramadhan ke-1	30	10	20
Ramadhan ke-2	42	14	28
Ramadhan ke-3	35	18	17
Ramadhan ke-4	39	12	27
Ramadhan ke-5	44	11	33
Ramadhan ke-6	40	17	23
Ramadhan ke-7	31	10	21
Ramadhan ke-8	43	18	25
Ramadhan ke-9	47	10	37
Ramadhan ke-10	37	9	28
Ramadhan ke-11	31	11	20
Ramadhan ke-12	40	15	25
Ramadhan ke-13	38	7	31
Ramadhan ke-14	42	12	30
Ramadhan ke-15	42	10	32
Ramadhan ke-16	35	7	28
Ramadhan ke-17	38	9	29
Ramadhan ke-18	33	11	22
Ramadhan ke-19	35	16	19
Ramadhan ke-20	26	6	20
Ramadhan ke-21	28	5	23
Ramadhan ke-22	31	5	26
Ramadhan ke-23	24	3	21
Ramadhan ke-24	20	7	13
Ramadhan ke-25	22	5	17
Ramadhan ke-26	28	5	23
Ramadhan ke-27	16	8	8
Ramadhan ke-28	20	8	12
Ramadhan ke-29	15	9	6
Ramadhan ke-30	12	4	8

2.4. Tahap Clustering

Proses pengelompokan data tanpa pemeliharaan dan memisahkan sekumpulan data dari himpunan sesuai kriteria menjadi beberapa kelas merupakan *Clustering*. Persamaan dan tahapan tentang jarak algoritma, atau biasa disebut Euclidean Distance dapat diterapkan dalam proses pengelompokan data (Adelina Hutasoit et al., 2019). Metode yang diimplementasikan untuk mengklasifikasikan kumpulan data menjadi beberapa bagian sesuai dengan kategori yang sama dimana telah ditetapkan sebelumnya ialah Kajian *cluster*. Dalam menentukan *cluster* berdasarkan data yang telah tersedia, dibutuhkan sebuah flowchart untuk memudahkan dalam menentukan alur perhitungan sebagai alur untuk menemukan hasil dari penerapan *cluster* terhadap data yang akan diproses. Berikut adalah flowchart dalam menentukan *cluster* dengan K-Means (Sianipar et al., 2020). Dapat dilihat pada Gambar 1.



Gambar 1. Diagram Alur K-Means

3. HASIL DAN PEMBAHASAN

Proses perhitungan Algoritma *K-Means* adalah proses penentuan nilai *cluster* dan pengelompokan data sehingga menghasilkan nilai-nilai *cluster* baru dari setiap pengelompokan. Pada kasus ini data responden pengunjung atau data sekunder sebagai Analisa. Proses ini Analisa responden pengunjung menggunakan metode Algoritma *K-Means* dapat dilihat sebagai berikut.

3.1. Tahap Perhitungan Clustering

Pada tahap ini data yang diambil dari data jumlah hari kegiatan (*id_pertanyaan*) sebanyak 30 data sebagai sampel. Terdapat: Jumlah Cluster : 3, Jumlah Data : 30, Jumlah Atribut : 3 (jumlah responden/hari, jumlah responden menjawab benar dan jumlah responden menjawab salah).

- Menentukan secara acak Nilai *k* sebagai jumlah *cluster* yang ingin dibentuk, tetapkan pusat *cluster*. Data yang dipakai adalah sebagai berikut.

$$M_1 = (n_1 40,538)(n_2 11,615)(n_3 28,923)$$

$$M_2 = (n_1 40,538)(n_2 11,615)(n_3 28,923)$$

$$M_3 = (n_1 30,182)(n_2 9,091)(n_3 21,091)$$

Menghitung centroid terdekat. Centroid terdekat yang digunakan yaitu Ramadhan ke-1 (30,10,20). Yaitu dilakukan perhitungan seperti dibawah ini. Diketahui :

$$DM_1 = \sqrt{\frac{(30 - 40,538)^2 + (10 - 11,615)^2}{+(20 - 28,923)^2}} = 13,90$$

$$DM_2 = \sqrt{\frac{(30 - 17,500)^2 + (10 - 6,833)^2}{+(20 - 10,667)^2}} = 15,92$$

$$DM_3 = \sqrt{\frac{(30 - 30,182)^2 + (10 - 9,091)^2}{+(21 - 28,091)^2}} = 1,43$$

Dan dilanjutkan sampai data ke-n

Kemudian didapatkan hasil perhitungan *cetroid* sampai data ke-n (iterasi ke-1). Dapat dilihat pada Tabel 2.

Tabel 2. Hasil Perhitungan Cetroid Pertama (iterasi ke-1)

Jumlah Respon	Jumlah respondent		C1	C2	C3
	benar	salah			
30	10	20	13,90	15,92	1,43
42	14	28	2,95	30,86	14,54
35	18	17	14,61	21,70	10,92
39	12	27	2,49	27,49	11,01
44	11	33	5,38	34,91	18,34
40	17	23	8,02	27,60	12,75
31	10	21	12,50	17,29	1,23
43	18	25	7,89	31,31	16,09
47	10	37	10,47	39,67	23,17
37	9	28	4,50	26,18	9,71
31	11	20	13,08	16,93	2,35
40	15	25	5,21	27,90	12,11
38	7	31	5,66	28,87	12,79
42	12	30	1,86	31,63	15,08
42	10	32	3,77	32,64	16,11
35	7	28	7,27	24,63	8,68
38	9	29	3,64	27,59	11,12
33	11	22	10,25	19,65	3,52
35	16	19	12,18	21,44	8,68
26	6	20	17,96	12,65	5,31
28	5	23	15,36	16,30	5,01
31	5	26	11,97	20,51	6,44
24	3	21	20,26	12,80	8,68
20	7	13	26,39	3,42	13,17
22	5	17	23,01	7,98	10,02
28	5	23	15,36	16,30	5,01
16	8	8	32,45	3,27	19,33
20	8	12	26,86	3,06	13,69
15	9	6	34,42	5,72	21,41
12	4	8	36,20	6,74	22,98

- c. Menghitung centroid terdekat. Centroid terdekat yang digunakan yaitu Ramadhan ke-1 (30,10,20). Yaitu dilakukan perhitungan seperti Tabel 3 dibawah ini.

Tabel 3. Pengelompokan Data Cluster (hitungan iterasi ke-1)

Cluster	ID_Pertanyaan
C1	Ramadhan ke 2
C1	Ramadhan ke 4
C1	Ramadhan ke 5
C1	Ramadhan ke 6
C1	Ramadhan ke 8
C1	Ramadhan ke 9
C1	Ramadhan ke 10
C1	Ramadhan ke 12
C1	Ramadhan ke 13
C1	Ramadhan ke 14
C1	Ramadhan ke 15
C1	Ramadhan ke 16
C1	Ramadhan ke 17
C2	Ramadhan ke 24
C2	Ramadhan ke 25
C2	Ramadhan ke 27
C2	Ramadhan ke 28
C2	Ramadhan ke 29
C2	Ramadhan ke 30
C3	Ramadhan ke 1
C3	Ramadhan ke 3
C3	Ramadhan ke 7
C3	Ramadhan ke 11
C3	Ramadhan ke 18
C3	Ramadhan ke 19
C3	Ramadhan ke 20
C3	Ramadhan ke 21
C3	Ramadhan ke 22
C3	Ramadhan ke 23
C3	Ramadhan ke 26

3.2. Membentuk Cluster Centroid Baru

- a. Membentuk *cluster* centroid baru dan untuk mendapatkan centroid baru yaitu mencari nilai rata-rata dari masing-masing *cluster*. Pada data *cluster* ke-1 memiliki 13 data, yaitu data ke-2,4,5,6,8,9,10,11,12,13,14,15,16,17.

$$C_{11} = \frac{42 + 39 + 44 + 40 + 43 + 47 + 37 + 40 + 38 + 42 + 42 + 35 + 38}{13} = 40,54$$

$$C_{12} = \frac{14 + 12 + 11 + 17 + 10 + 18 + 10 + 9 + 15 + 12 + 10 + 7 + 9}{13} = 11,62$$

$$C_{13} = \frac{28 + 27 + 33 + 23 + 25 + 37 + 28 + 25 + 30 + 32 + 28 + 29 + 25}{13} = 28,923$$

- b. Pada *cluster* ke-2 memiliki 6 data, yaitu data ke-24,25,27,28,29,30.

$$C_{21} = \frac{20 + 22 + 16 + 20 + 15 + 12}{6} = 17,50$$

$$C_{22} = \frac{7 + 5 + 8 + 8 + 9 + 4}{6} = 6,83$$

$$C_{23} = \frac{13 + 17 + 8 + 12 + 6 + 8}{6} = 10,67$$

c. Pada *cluster* ke-3 memiliki 11 data, yaitu data ke-1,3,7,11,18,19,20,22,23,26.

$$C_{31} = \frac{30 + 35 + 31 + 31 + 33 + 35 + 26 + 28 + 31 + 24 + 28}{11} = 30,18$$

$$C_{32} = \frac{10 + 18 + 10 + 11 + 11 + 16 + 6 + 5 + 5 + 3 + 5}{11} = 9,09$$

$$C_{33} = \frac{20 + 17 + 21 + 20 + 22 + 19 + 20 + 23 + 26 + 21 + 23}{11} = 21,09$$

d. Setelah perhitungan centroid baru yaitu mendapatkan hasil nilai.

$$M_1 = (n_1 40,54)(n_2 11,62)(n_3 28,923)$$

$$M_2 = (n_1 17,50)(n_2 6,83)(n_3 10,67)$$

$$M_3 = (n_1 30,182)(n_2 9,091)(n_3 21,091)$$

Dilanjutkan menghitung iterasi ke-2

e. Setelah mendapatkan hasil nilai cetroid baru, yaitu kembali pada langkah perhitungan ke 2 dan iterasi 1 serta melakukan perhitungan Kembali untuk mendapatkan hasil perhitungan jarak baru dan hasil setelah perhitungan tersebut Dapat dilihat pada Tabel 4.

Tabel 4. Hasil Perhitungan Cetroid Pertama (iterasi ke-2)

Jumlah Respon	Jumlah respondent		C1	C2	C3
	benar	salah			
30	10	20	13,90	15,92	1,43
42	14	28	2,95	30,86	14,54
35	18	17	14,61	21,70	10,92
39	12	27	2,49	27,49	11,01
44	11	33	5,38	34,91	18,34
40	17	23	8,02	27,60	12,75
31	10	21	12,50	17,29	1,23
43	18	25	7,89	31,31	16,09
47	10	37	10,47	39,67	23,17
37	9	28	4,50	26,18	9,71
31	11	20	13,08	16,93	2,35
40	15	25	5,21	27,90	12,11
38	7	31	5,66	28,87	12,79
42	12	30	1,86	31,63	15,08
42	10	32	3,77	32,64	16,11
35	7	28	7,27	24,63	8,68
38	9	29	3,64	27,59	11,12
33	11	22	10,25	19,65	3,52
35	16	19	12,18	21,44	8,68
26	6	20	17,96	12,65	5,31
28	5	23	15,36	16,30	5,01

31	5	26	11,97	20,51	6,44
24	3	21	20,26	12,80	8,68
20	7	13	26,39	3,42	13,17
22	5	17	23,01	7,98	10,02
28	5	23	15,36	16,30	5,01
16	8	8	32,45	3,27	19,33
20	8	12	26,86	3,06	13,69
15	9	6	34,42	5,72	21,41
12	4	8	36,20	6,74	22,98

- f. Dan hasil iterasi ke-1 dan ke-2 tidak ada perubahan posisi nilai *cluster* yaitu proses iterasi selanjutnya diberhentikan dan pada hasil akhirnya mendapatkan 3 buah *cluster*. Dapat dilihat pada Tabel 5.

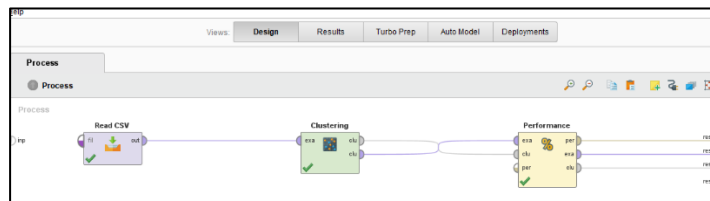
Tabel 5. Pengelompokan Data Cluster (hitungan iterasi ke-1)

Cluster	ID_Pertanyaan
C1	Ramadhan ke 2
C1	Ramadhan ke 4
C1	Ramadhan ke 5
C1	Ramadhan ke 6
C1	Ramadhan ke 8
C1	Ramadhan ke 9
C1	Ramadhan ke 10
C1	Ramadhan ke 12
C1	Ramadhan ke 13
C1	Ramadhan ke 14
C1	Ramadhan ke 15
C1	Ramadhan ke 16
C1	Ramadhan ke 17
C2	Ramadhan ke 24
C2	Ramadhan ke 25
C2	Ramadhan ke 27
C2	Ramadhan ke 28
C2	Ramadhan ke 29
C2	Ramadhan ke 30
C3	Ramadhan ke 1
C3	Ramadhan ke 3
C3	Ramadhan ke 7
C3	Ramadhan ke 11
C3	Ramadhan ke 18
C3	Ramadhan ke 19
C3	Ramadhan ke 20
C3	Ramadhan ke 21
C3	Ramadhan ke 22
C3	Ramadhan ke 23
C3	Ramadhan ke 26

3.3. Tahap Pengujian Algoritma K-Means Pada Rapid Miner

Pada proses pengujian *K-Means* pada RepidMiner ini menggunakan data yang sama tanpa ada perubahan nilai dan jumlah *cluster* berikut adalah hasil pemrosesan data pada Repid Miner.

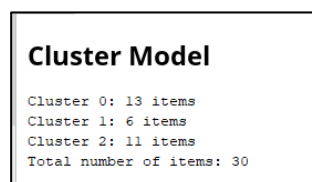
Desain proses : merupakan tampilan desain permodelan *k-means clustering* dilihat pada Gambar 2.



Gambar 2. Permodelan Clustering K-Means pada Rapid Miner

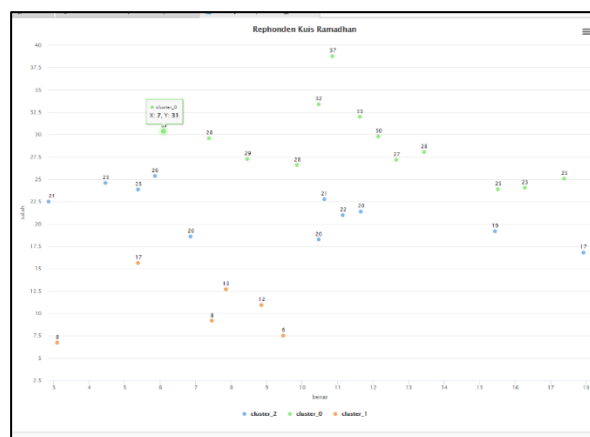
Pada hasil pemrosesan permodelak *K-Means clustering* seperti pada Gambar 2, dengan inisialisai jumlah *cluster* model ada 3, dengan nilai *cluster* 0 = 13, *cluster* 1 = 6, *cluster* 2 = 11, dengan jumlah data 30.

Text View : merupakan tampilan sheet untuk data base yang telah diolah secara keseluruhan berdasarkan *clusternya*. Dilihat pada Gambar 3.



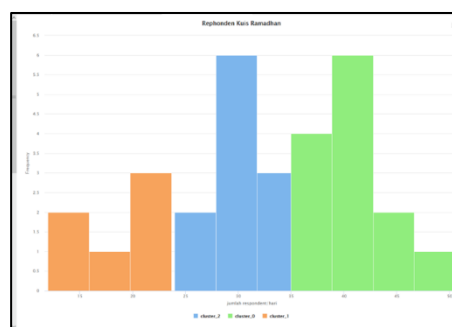
Gambar 3. Hasil Data Clustering K-Means Pada Rapid Miner

Flot View : dimana visualisasi hasil pengelompokan jumlah reshponden menjawab benar dan jumlah reshponden menjaawab salah berdasarkan pengelompokan *cluster*. Dilihat pada Gambar 4.



Gambar 4. Bubble Nilai *clastering* jumlah reshponden menjawab benar & salah berdasarkan *cluster*

Visualisasi Histogram clutering hasil pengelompokan data jumlah reshponden berdasarkan pengelompokan *clustering*, digambarkan pada Gambar 5.



Gambar 5. Histogram clustering Jumlah Reshponden pada pengelompokan cluster

4. KESIMPULAN

Pada penelitian ini penulis memiliki kesimpulan bahwa algoritma *K-Means* dapat digunakan sebagai analisis data pengunjung mall untuk mengetahui jumlah reshpondent pada suatu event dan aplikasi RapidMiner memudahkan untuk melihat Hasil secara visualisai pada pengelompokan data *clustering*. Dari semua data yang telah diolah dapat disimpulkan bahwa penelitian ini, penulis mengelompokan data pengunjung yang dimana jumlah data responden/hari, jumlah reshponden menjawab benar dan jumlah reshponden menjawab benar, yaitu jumlah yang *cluster* yang digunakan yaitu 3 *cluster* dimana nilai cluter tersebut yaitu : *cluster_0*, *cluster_1*, *cluster_3*. Selain itu, berdasarkan hasil pengujian dari algoritma *K-Means*, medapatkan *cluster* model untuk algoritma *K-Means* sebanyak 13 item pada *cluster_0*, atau jumlah reshponden tinggi, *cluster_1* sebanyak 6 item jumlah reshpondent rendah, dan *cluster_2* sebanyak 11 dengan reshpondent sedang. Adapun penelitian ini juga memvisualisasikan data *K-Means* pada RapidMiner dengan hasil pengelompokan data jumlah reshpondent berdasarkan *cluster* dan histogram hasil pengelompokan jumlah reshpondent/hari berdasarkan *cluster*.

DAFTAR PUSTAKA

- Adelina Hutasoit, R., Parlina, I., Tunas Bangsa Pematangsiantar, S., Tunas Bangsa Pematangsiantar, A., & Jendral Sudirman Blok No, J. A. (2019). *Prosiding Seminar Nasional Riset Information Science (SENARIS) Penerapan Data Mining Dalam Mengelompokkan Jumlah Usaha Berdasarkan Provinsi Menggunakan K-Means Clustering*. September, 937–945. <http://www.bps.go.id>.
- Anggreini, N. L. (2019). Teknik Clustering Dengan Algoritma K-Medoids Untuk Menangani Strategi Promosi Di Politeknik Tedc Bandung. *Jurnal Teknologi Informasi Dan Pendidikan*, 12(2), 1–7. <https://doi.org/10.24036/tip.v12i2.215>
- Dimas Hendika Wibowo. (2015). ANALISIS STRATEGI PEMASARAN UNTUK MENINGKATKAN DAYA SAING UMKM (Studi pada Batik Diajeng Solo). *Jurnal Administrasi Bisnis (JAB)*, 29(1), 59–66.
- Purnamasari, S. M. (2011). Analisis Kelompok (Cluster Analysis). *Program*, 18209007, 1–6. <http://informatika.stei.itb.ac.id/~rinaldi.munir/Probststat/2010-2011/Makalah2010/MakalahProbststat2010-030.pdf>
- Sianipar, K. D. R., Siahaan, S. W., Siregar, M., R.H Zer, F. I., & Hartama, D. (2020). Penerapan Algoritma K-Means Dalam Menentukan Tingkat Kepuasan Pembelajaran Online Pada Masa Pandemi Covid-19. *Jurnal Teknologi Informasi*, 4(1), 101–105. <https://doi.org/10.36294/jurti.v4i1.1258>
- Sugianto, C. A., Rahayu, A. H., & Gusman, A. (2020). Algoritma K-Means untuk Pengelompokkan Penyakit Pasien pada Puskesmas Cigugur Tengah. *Journal of Information Technology*, 2(2), 39–44. <https://doi.org/10.47292/joint.v2i2.30>
- Triyansyah, D., & Fitriana, D. (2018). Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing. *Jurnal Telekomunikasi Dan Komputer*, 8(3), 163. <https://doi.org/10.22441/incomtech.v8i3.4174>
- Willy Pratama Widharta dan Sugiono Sugiharto, S.E., M. . (2012). PENYUSUNAN STRATEGI DAN SISTEM PENJUALAN DALAM RANGKA MENINGKATKAN PENJUALAN TOKO DAMAI. *Dictionary of Marketing Communications*, 2(1), 1–15. <https://doi.org/10.4135/9781452229669.n3693>

Halaman Ini Dikosongkan