

Algoritma K-Means Clustering Penggunaan Bandwidth Internet (Studi Kasus di Pemerintah Daerah Kabupaten Padang Pariaman)

Rizki Mubarak , Sarjon Defit, Gunadi Widi Nurcahyo

Jurusan Teknik Informatika

Fakultas Ilmu Komputer

Universitas Putra Indonesia YPTK Padang

Sumatera Barat, Indonesia

rizkimubarak88@gmail.com, sarjond@yahoo.co.uk, gunadiwidi@yahoo.com

Abstract-To support government activities, a fast and precise network connection is needed. So it requires a wide network bandwidth. Bandwidth management needs to be done so that the network speed remains stable. This study aims to look at the pattern of bandwidth usage in the Regional Government of Padang Pariaman Regency using K-Means Clustering. The data is taken from the Cacti application, an open-source, web-based network monitoring software. The total extracted datasets used are 32 OPD data (Regional Apparatus Organizations) in the Regional Government of Padang Pariaman Regency in 2022. The available data is then processed to obtain cluster targets by utilizing the data mining concept using the K-Mean Clustering method. Bandwidth usage data grouping in Padang Pariaman Regency uses the Clustering method with the K-Means algorithm with the attributes Name OPD, Inbound Average, Inbound Maximum, Outbound Average, Outbound Maximum used in the process of calculating and dividing data into 3 clusters with high bandwidth usage categories, low, and medium. Calculations are done manually and then tested with RapidMiner software. The results of the manual calculations obtained the same number of cluster members as the calculations with the RapidMiner software.

Keywords: Breadth First Search, Certainty Factor, Coffee, Expert System.

Abstrak-Untuk menunjang kegiatan di Pemerintahan dibutuhkan koneksi jaringan yang yang cepat dan tepat. Sehingga memerlukan jaringan bandwidth yang lebar. Manajemen Bandwidth perlu dilakukan agar kecepatan jaringan tetap stabil. Penelitian ini bertujuan untuk melihat pola penggunaan bandwidth di Pemerintah Daerah Kabupaten Padang Pariaman menggunakan K-Means Clustering. Data diambil dari aplikasi Cacti sebuah software open-source, pemantauan jaringan berbasis web. Total datasets hasil ekstraksi yang digunakan adalah sebanyak 32 data OPD (Organisasi Perangkat Daerah) yang ada di Pemerintah Daerah Kabupaten Padang Pariaman tahun 2022.. Data-data yang tersedia selanjutnya diolah untuk mendapatkan target cluster dengan memanfaatkan konsep data mining menggunakan metode K-Mean Clustering. Pengelompokan data penggunaan bandwidth di Kabupaten Padang Pariaman menggunakan metode Clustering dengan algoritma K-Means dengan atribut Nama OPD, Inbound Average, Inbound Maksimum, Outbound Average, Outbound Maximum yang digunakan dalam proses perhitungan dan pembagian data ke dalam 3 cluster dengan kategori penggunaan bandwidth tinggi, rendah, dan sedang. Perhitungan dilakukan secara manual dan kemudian dilakukan pengujian dengan software RapidMiner. Hasil dari perhitungan manual diperoleh jumlah anggota cluster yang sama dengan perhitungan dengan software RapidMiner.

Kata Kunci: K-Means Algorithm, Clustering, Bandwidth Usage, Data Mining

1. Pendahuluan

Manajemen bandwidth sangat diperlukan agar tidak terjadi pemborosan penggunaan bandwidth, tersedia konektivitas yang stabil, dan membantu administrator jaringan untuk mengontrol penggunaan bandwidth. Walaupun, sistem administrasi jaringan mampu mendapatkan data penggunaan menggunakan bandwidth dengan menggunakan aplikasi Cacti. Namun, aplikasi tersebut belum dapat digunakan dalam menganalisis penggunaan bandwidth dengan tepat. Karena itu diperlukan metode manajemen bandwidth analisis data. Manajemen bandwidth penting untuk dilakukan agar dapat

mengatur bandwidth secara merata kepada seluruh pengguna [1].

Knowledge Discovery Database (KDD) adalah penerapan metode saintifik pada data mining. Penemuan Knowledge Discovery Database (KDD) muncul dan menjadi hotspot penelitian di bidang kecerdasan buatan dan aplikasi database dalam beberapa tahun terakhir. Saat ini, penelitian KDD mencakup berbagai macam metode penemuan pengetahuan di banyak bidang dan telah mampu menemukan berbagai jenis pengetahuan seperti aturan deret waktu, aturan asosiasi, aturan klasifikasi, dan aturan clustering. Penemuan pengetahuan dalam basis

Vol.14 no.1 | Juni 2023

EXPLORE : ISSN: 2087-2062, Online ISSN: 2686-181X / DOI: <https://dx.doi.org/10.36448/jsit.v14i1.3037>



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

data (KDD) adalah proses penentuan informasi dan pola yang berguna dalam data [2].

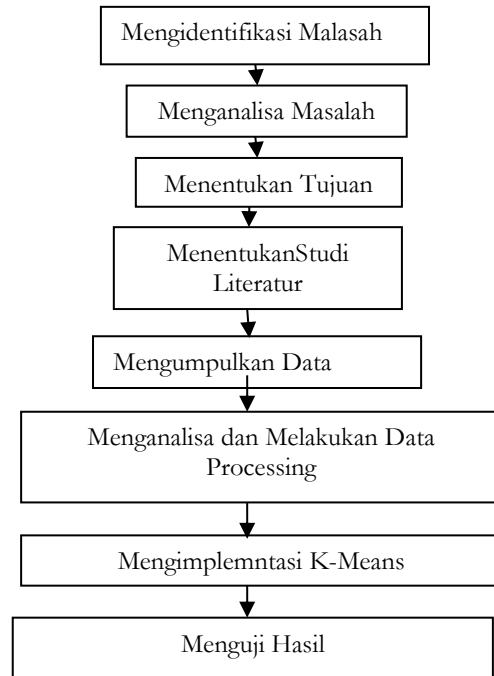
Pendekatan K-means, salah satu metode pengelompokan yang paling banyak digunakan, untuk mendapatkan kluster ini [3]. Kualitas segmentasi dan efisiensi algoritma bergantung pada pemilihan centroid awal, jumlah cluster yang dipilih, dan tingkat iterasi yang akan dilakukan[4]. Algoritma K-means clustering lebih efektif dan efisien dalam mengolah data dengan jumlah yang banyak [5]. K-Means adalah suatu metode yang digunakan untuk membagi suatu data kedalam beberapa kelompok atau cluster [6]. Clustering (berbasis centroid) adalah algoritma clustering yang menggunakan titik centroid dalam melakukan pengelompokan objek [7]. K-means merupakan algoritma sederhana dengan waktu proses yang cepat dan menghasilkan cluster yang optima [8]. K-Means adalah algoritma iteratif; itu terus mengulangi langkah-langkah di atas sampai tidak ada perubahan lokasi centroid [9], Tujuan dari metode clustering data ini adalah untuk meminimalkan fungsi tujuan yang ditetapkan dalam proses clustering [10]. Algoritma K-means memberikan metode sederhana untuk mengeksekusi solusi perkiraan[11]. Ini adalah pengelompokan eksklusif dan salah satu algoritma yang paling banyak digunakan untuk pengelompokan [12]. Algoritma ini sudah banyak digunakan pada penelitian sebelumnya[13]. kelebihan algoritma K-Means yaitu mampu mengelompokkan objek dalam jumlah besar dengan pengelompokan yang cepat, dan kekurangan dari algoritma K-Means yaitu hasil pengelompokannya yang mudah berubah-ubah.[14]. Kekuatan algoritma K-means adalah karena efisiensi komputasi dan sifat kemudahan yang dapatdigunakan [15]. Salah satu algoritma untuk analisis cluster adalah K-Means clustering [16]. Namun, K-means mempunyai kelemahan yang diakibatkan oleh penentuan pusat awal cluster. Hal ini menyebabkan hasil cluster-nya berupa solusi yang sifatnya local optimal. [17]. Algoritma k-Means clustering digunakan untuk mengelompokkan [18] atau mempartisi sejumlah data n menjadi sebanyak k cluster yang mana tiap data tersebut masuk ke suatu cluster dengan rerata terdekat. K-Means yang merupakan algoritma pengelompokkan iterative yang sederhana diimplementasikan[19]

Penelitian terdahulu yang berkaitan dengan Algoritma K-Means Clustering telah banyak dilakukan, seperti pada penelitian yang dilakukan oleh Andi Widya Mufila Gaffaradan yang membahas mengenai Pengklasifikasian Kecepatan Transfer Data Pada Jaringan Backbone Menggunakan K-Means[20]. Hasil penelitian terdapat tiga cluster yaitu kecepatan transfer yang rendah, normal dan kecepatan transfer yang tinggi. Penelitian terkait lainnya dilakukan oleh Sasa Ani Arnomo dan Yulia, yang menerapkan K-Means Clustering mengenai pengelompokan potensi peningkatan bandwidth FTTH pelanggan pita lebar. Tujuan penelitian ini yaitu menentukan pelanggan yang berpotensi untuk dilakukan penambahan bandwidth. Penelitian lainnya adalah penelitian yang dilakukan mengenai analisis data lalu lintas jaringan di kantor Cangehgar Cyber Operation Center menggunakan algoritma K-Means [21]. Penelitian

bertujuan untuk menganalisa lalu lintas jaringan di Kantor Cangehgar Cyber Operation Center.

2. Metodologi

Adapun tahapan penelitian yang dilakukan dapat dilihat pada Gambar 1



Gambar 1. Tahapan Penelitian

A. Pengumpulan Data

Data yang digunakan dalam penelitian adalah penggunaan bandwidth di beberapa dinas di Kabupaten Padang Pariaman. Data yang digunakan sebanyak 32 sampel data bandwidth dari total sebanyak 55 data . Data diambil dengan cara meminta kepada Admin Jaringan dan mewawancarai admin jaringan mengenai alokasi penggunaan bandwidth, sehingga penulis dapat mendapatkan informasi beserta data untuk perhitungan menggunakan metoda K-Means. Data meliputi 4 kriteria sebagai berikut :

1. Inbound Average tahun 2022.
2. Inbound Maksimum tahun 2022.
3. Outbound Average tahun 2022
4. Outbound Maximum tahun 2022

B. Perhitungan dengan K-Means Clustering

Pada tahap ini akan dilakukan pengolahan data *Bandwidth* menggunakan metode K-Means *Clustering* hingga didapatkan hasil kluster untuk setiap *bandwidth*. Berikut langkah – langkah penyelesaiannya

1. Pilih secara acak k buah data sebagai pusat *cluster*
2. 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 yang dirumuskan pada Persamaan (1)



$$dist(x, y) = \sqrt{\sum_{i=1}^n (X_i - X_y)^2} \quad : i = 1, 2, 3, \dots, n \quad (1)$$

3. Data ditempatkan dalam *cluster* yang terdekat, dihitung dari tengah *cluster*
4. Pusat *cluster* baru akan ditentukan bila semua data telah ditetapkan dalam *cluster* terdekat
5. Proses penentuan pusat *cluster* dan penempatan data dalam *cluster* diulangi sampai nilai centroid tidak berubah lagi.

Data yang diperoleh peneliti menggunakan 4 atribut. Berikut adalah data yang digunakan yaitu data pemakaian *Bandwidth* di Pemerintah Daerah Kabupaten Padang Pariaman tahun 2022.

Tabel 1 Data Penggunaan Bandwidth

OPD	Inbound Average	Inbound Maksimum	Outbound Average	Outbound Maksimum
Setwan	4.26	9.14	0.343	1.46
Inspektorat	1.59	4.37	0.178	1.28
Disdikbud	2.82	8.57	0.215	1.2
Dinkes	4.67	10.69	0.546	3.24
DinsoP3A	3.32	5.97	0.281	0.939
DLHKP	2.61	7.63	0.227	1.89
DPMD	3.48	8.66	0.35	1.61
DPPKB	2.4	7.36	0.222	1.81
Disnagnerko P	4.12	10.35	0.532	3.22
DPMPTP	2.06	7.01	0.202	0.992
Dinas Pariwisata	2.28	7.76	0.325	1.67
Dinas Kearsipan	1.48	8.24	0.166	0.844
Dinas Pertanian	2.53	6.91	0.329	2.14
Dinas Peternakan	2.71	6.3	0.236	1.33
Dinas PUPR	4.23	12.3	0.433	1.66
Dinas Perikanan	1.63	5.14	0.156	1.15
Disdukcapil	5.74	11.63	0.704	4.77
BPBD	7.11	12.62	0.64	2.94
Bappeda-BKPSDM	5.89	15.1	0.552	4.1
Kec. Batang Anai	2.84	8.59	0.216	2.09
Kec. Nan Sabaris	1.42	3.19	0.82	0.162

Kec. 2X11 Kayutanam	2.52	6.89	0.185	0.65
Kec. V Koto Dalam	2.15	5.33	0.209	0.171
Kec. Ulakan Tapakis	2.86	6.19	0.203	1.03
Kec. Sintoga	1.29	3.7	0.112	0.829
Kec. Padang Sago	3.65	3.65	0.26	0.26
Batang Gasan	1.39	4.13	0.127	1.27
Kec. V Koto Timur	2.27	5.09	0.185	1.02
Kec 2x11 Enam Lingkung	1.04	2.74	0.1	0.287
Kec Patamuam	2.33	4.67	0.482	1.63
Kec. Enam Lingung	2.5	4.85	0.189	0.951
Kec IV Koto Aur	0.82	3.14	0.061	0.354
Malintang	3			

Proses perhitungan yang peneliti adalah menggunakan *Microsoft Excel* dengan langkah – langkah berikut:

1. Menentukan jumlah cluster

Jumlah *cluster* merupakan jumlah kelompok yang akan dihasilkan. Jumlah ini menyesuaikan dengan kebutuhan analisis dalam penelitian ini jumlah cluster yang akan digunakan sebanyak 3 *cluster*.

2. Menentukan Centroid Awal

Proses awal *clustering* algoritma k-means adalah menentukan titik pusat awal *cluster (centroid)*. *Centroid* awal ditentukan secara acak seperti tabel 2 dibawah

Tabel 2 Centroid Awal

	Inbound Average	Inbound Maksimum	Outbound Average	Outbound Maksimum
C 1	2.4	7.36	0.222	1.81
C 2	2.53	6.91	0.329	2.14
C 3	5.89	15.1	0.552	4.1

3. Menghitung Jarak Data ke Titik Pusat Cluster

Langkah selanjutnya adalah menghitung jarak setiap data dengan *centroid* awal menggunakan persamaan *Euclidean Distance*. Pada tahap ini, jarak terdekat antara data dan *cluster* akan menentukan data mana yang termasuk dalam kelompok data mana. Berikut adalah hasil perhitungan jarak pada iterasi ke-1.

Tabel 3 Hasil Iterasi ke-1

OPD	Jarak ke Centroid			Kelompok
	C1	C2	C3	



	2.6	2.9	6.72	C1
Setwan	01	03	2	
	3.1	2.8	11.9	C2
Inspektorat	43	46	04	
	1.4	1.9	7.78	C1
Disdikbud	19	33	4	
	4.2	4.4	4.65	C1
Dinkes	89	86	6	
	1.8	1.7	10.0	C2
DinsoP3A	82	18	01	
	0.3	0.7	8.45	C1
DLHKP	51	73	9	
	1.7	2.0	7.31	C1
DPMD	07	61	6	
	0.0	0.5	8.80	C1
DPPKB	00	83	0	
	3.7	3.9	5.14	C1
DISDAGNAKER	39	46	5	
	0.9	1.2	9.48	C1
DPMPTP	53	51	2	
	0.4	1.0	8.53	C1
Dinas Pariwisata	52	03	6	
	1.5	2.1	8.79	C1
Dinas Kearsipan	99	40	0	
	0.5	0.0	9.07	C2
Dinas Pertanian	83	00	0	
	1.2	1.0	9.76	C2
Dinas Peternakan	04	34	3	
	5.2	5.6	4.07	C3
Dinas PUPR	74	73	0	
	2.4	2.2	11.2	C2
Dinas Perikanan	42	26	34	
	6.1	6.2	3.54	C3
Disdukcapil	95	96	1	
	7.1	7.3	2.99	C3
BPBD	63	70	9	
	8.8	9.0	0.00	C3
Bappeda-BKPSDM	00	70	0	
	1.3	1.7	7.47	C1
Kec. Batang Anai	36	13	2	
	4.6	4.3	13.3	C2
Kec. Nan Sabaris	28	85	19	
Kec. 2X11	1.2	1.4	9.52	C1
Kayutanam	58	97	9	
Kec. V Koto	2.6	2.5	11.1	C2
Kampung Dalam	21	56	80	
Kec. Ulakan	1.4	1.3	9.90	C2
Tapakis	80	69	5	
	3.9	3.6	12.7	C2
Kec. Sintoga	50	89	28	
	4.2	3.9	12.2	C2
Kec. Padang Sago	11	27	86	
	3.4	3.1	12.1	C2
Batang Gasan	28	35	98	
	2.4	2.1	11.0	C2
Kec. V Koto Timur	07	58	87	
Kec 2x11 Enam	5.0	4.8	13.8	C2
Lingkung	53	06	22	
	2.7	2.3	11.2	C2
Kec Patamaian	09	11	94	

	2.6	2.3	11.2	C2
Kec. Enam Lingung	55	83	52	
Kc IV Koto Aur	4.7	4.5	13.5	C2
Malintang	37	15	27	

4. Menghitung Titik Pusat Centroid Baru

Pada iterasi pertama data telah menghasilkan keanggotaan *cluster*, selanjutnya penulis melakukan perhitungan untuk menentukan nilai titik pusat *cluster* baru yang diperoleh dari menghitung rata-rata keanggotaan dari *cluster* 1, *cluster* 2 dan *cluster* 3, berikut merupakan nilai *centroid* baru:

Tabel 4 Titik Pusat *Cluster* Iterasi ke-2

	Inbound Average	Inbound Maksimum	Outbound Average	Outbound Maksimum
C1	2.962	8.408	0.294	1.723
C2	2.094	4.711	0.246	0.925
C3	5.743	12.913	0.582	3.368

Langkah selanjutnya adalah menghitung jarak setiap data dengan *centroid* baru menggunakan persamaan *Euclidean Distance*. Pada tahap ini, jarak terdekat antara data dan *cluster* akan menentukan data mana yang termasuk dalam kelompok data mana. Berikut adalah hasil perhitungan jarak pada iterasi ke-2.

Tabel 5 Hasil Iterasi ke-2

OPD	Jarak ke <i>Centroid</i>			Kelompok
	C1	C2	C3	
Setwan	1.514	4.960	4.486	C1
	52551	51613	12965	
Inspektorat	5	3	8	C2
	4.288	0.707	9.733	
Disdikbud	65946	40833	37746	C1
	8	7	2	
Dinkes	0.571	3.936	5.677	C3
	19074	80000	25649	
DinsoP3A	6	8	5	C2
	3.239	6.916	2.471	
DLHKP	28387	47891	30184	C1
	7	3	6	
DPMD	2.585	1.758	7.749	C1
	46634	04029	52749	
Dinkes	1	5		C2
	0.872	3.117	6.326	
DinsoP3A	10370	73576	65561	C1
	3	6	8	
DPMD	0.590	4.242	5.132	C1
	18757	47811	77788	
		5	5	



DPPKB	1.193 93915 3	2.810 03693 1	6.675 18904 7	C1
DISDAGNAKER	2.722 61927 2	6.423 06206 6	3.036 97115 8	C1
DPMPTP	1.819 02429	2.301 00702	7.361 24329 3	C1
Dinas Pariwisata	0.942 17330 4	3.145 53223 7	6.440 87310 2	C1
Dinas Kearsipan	1.735 63783 7	3.584 17616 8	6.822 20483 5	C1
Dinas Pertanian	1.613 67609 3	2.551 4981	6.922 50347 1	C1
Dinas Peternakan	2.159 33208 1	1.752 03005 2	7.562 57283	C2
Dinas PUPR	4.096 76610 7	7.920 63643 6	2.366 56804 5	C3
Dinas Perikanan	3.577 32978 2	0.676 86911 9	9.078 73382 2	C2
Disdukcapil	5.249 36977 5	8.727 21763	1.904 37438 9	C3
BPBD	6.046 03779 6	9.588 23213 9	1.463 45611 9	C3
Bappeda-BKPSDM	7.686 43886 1	11.51 18452 1	2.311 79233 8	C3
Kec. Batang Anai	0.434 62056 4	4.118 71558 3	5.373 51447 5	C1
Kec. Nan Sabaris	5.684 39545 6	1.918 07131 1	11.11 49794 3	C2
Kec. 2X11 Kayutanam	1.913 40112 6	2.238 43540 3	7.361 90371 5	C1
Kec. V Koto Kampung Dalam	3.541 99833 4	0.978 21236 4	8.986 50601 2	C2
Kec. Ulakan Tapakis	2.327 27040 4	1.669 78862 6	7.688 20845 9	C2
Kec. Sintoga	5.078 13115 6	1.301 82667	10.55 27261 8	C2
Kec. Padang Sago	5.024 85220 7	1.997 23322 5	9.996 64262 7	C2
Batang Gasan	4.582 60495 8	0.982 64505 5	10.03 41079	C2

Kec. V Koto Timur	3.462 70374 1	0.433 09705 5	8.883 59872	C2
Kec 2x11 Enam Lingkung	6.157 36334 5	2.328 64782 2	11.63 25108 1	C2
Kec Patamuan	3.796 29697 2	0.781 06901 4	9.089 16491 3	C2
Kec. Enam Lingkung	3.670 96293 5	0.433 78644 5	9.028 39023 9	C2
Kc IV Koto Aur Malintang	5.852 25941	2.107 70975 9	11.36 02869 8	C2

5. Menghitung Titik Pusat Centroid Baru (Iterasi 3)

Pada iterasi pertama data telah menghasilkan keanggotaan *cluster*, selanjutnya penulis melakukan perhitungan untuk menentukan nilai titik pusat *cluster* baru yang diperoleh dari menghitung rata-rata keanggotaan dari *cluster* 1, *cluster* 1, dan *cluster* 3, berikut merupakan nilai *centroid* baru:

Tabel 6 Titik Pusat *Cluster* Iterasi ke-3

	Inbound Average	Inbound Maksimum	Outbound Average	Outbound Maksimum
C1	2.783	8.093	0.276	1.631
C2	2.065	4.564	0.240	0.844
C3	5.528	12.468	0.575	3.342

Langkah selanjutnya adalah menghitung jarak setiap data dengan centroid baru menggunakan persamaan Euclidean Distance. Pada tahap ini, jarak terdekat antara data dan cluster akan menentukan data mana yang termasuk dalam kelompok data mana. Berikut adalah hasil perhitungan jarak pada iterasi ke-2.

Tabel 7 Hasil Iterasi ke-3

OPD	Jarak ke <i>Centroid</i>			Kelompok
	C1	C2	C3	
Setwan	1.819 79251 7	5.113 53290 1	4.034 74361	C1
Inspektorat	3.926 07817 1	0.675 93766 6	9.246 34527 8	C2
Disdikbud	0.647 39488 6	4.092 12385 7	5.219 71570 1	C1
Dinkes	3.601 00344 4	7.081 53444 7	1.977 04147 7	C3



DinsoP3A	2.296	1.887	7.277	C2
	16371	55640	36992	
	5	9	3	
DLHKP	0.559	3.285	5.843	C1
	69647	02530	83230	
		6	4	
DPMD	0.901	4.402	4.663	C1
	84910	08844	21959	
	4	3	6	
DPPKB	0.847	2.977	6.192	C1
	54850	08273	55044	
	7	2	4	
DISDAGNAKER	3.077	6.590	2.546	C3
	72347	22743	59007	
	8	1	3	
DPMPTP	1.452	2.450	6.890	C1
	32244	75977	45840	
	1	2	3	
Dinas Pariwisata	0.606	3.309	5.964	C1
	46253	06033	29811	
	4	9	5	
Dinas Kearsipan	1.533	3.722	6.377	C1
	76263	97074	27002	
		9	7	
Dinas Pertanian	1.313	2.721	6.433	C1
	02505	59797	09319	
	9	3	1	
Dinas Peternakan	1.819	1.914	7.081	C1
	57022	65666	55442	
	6	1		
Dinas PUPR	4.452	8.076	2.135	C3
	11988	89932	95786	
	5	6	5	
Dinas Perikanan	3.208	0.788	8.595	C2
	35094	31685	02257	
	8	4	1	
Disdukcapil	5.593	8.891	1.674	C3
	76407	69413	22011	
		2	7	
BPBD	6.408	9.741	1.640	C3
	07343	91651	62701	
	4	2	4	
Bappeda-BKPSDM	8.057	11.67	2.762	C3
	72456	63247	88997	
	3	8	2	
Kec. Batang Anai	0.681	4.285	4.894	C1
	68355	10188	95791	
	7	1	6	
Kec. Nan Sabaris	5.324	1.762	10.63	C2
	28966	27216	62292	
	8	8	7	
Kec. 2X11 Kayutanam	1.576	2.378	6.896	C1
	91052	68738	45648	
	6	2	1	
Kec. V Koto Kampung Dalam	3.188	1.023	8.517	C2
	97786	79626	69481	
	2	7	7	
Kec. Ulakan Tapakis	1.998	1.819	7.212	C2
	07777	88975	1589	
	8	7		

Kec. Sintoga	4.711	1.167	10.06	C2
	12889	69530	81679	
	6	7	6	
Kec. Padang Sago	4.729	1.920	9.533	C2
	45329	86855	20287	
	5	2	2	
Batang Gasan	4.218	0.915	9.546	C2
	47683	34725	68403	
	5	4	2	
Kec. V Koto Timur	3.108	0.593	8.401	C2
	13880	86782	97786	
	1	7	2	
Kec 2x11 Enam Lingkung	5.790	2.169	11.15	C2
	22240	64994	05505	
	3		7	
Kec Patamuan	3.458	0.870	8.600	C2
	53366	40771	90698	
	9	8	7	
Kec. Enam Lingung	3.326	0.533	8.548	C2
	33529	98264	01643	
	2	2	7	
Kc IV Koto Aur Malintang	5.481	1.960	10.87	C2
	60272	18417	84626	
	8	4	2	

6. Menghitung Titik Pusat Centroid Baru (Iterasi 4)

Pada iterasi pertama data telah menghasilkan keanggotaan *cluster*, selanjutnya penulis melakukan perhitungan untuk menentukan nilai titik pusat cluster baru yang diperoleh dari menghitung rata-rata keanggotaan dari cluster 1, cluster 2, cluster 3 dan cluster 3, berikut merupakan nilai centroid baru:

Tabel 8 Titik Pusat *Cluster* Iterasi ke-4

	Inbound AVERAGE	Inbound Maksimum	Outbound Average	Outbound Maksimum
C1	2.666	7.755	0.251	1.474
C2	2.019	4.440	0.240	0.810
C3	5.293	12.115	0.568	3.322

Langkah selanjutnya adalah menghitung jarak setiap data dengan centroid baru menggunakan persamaan Euclidean Distance. Pada tahap ini, jarak terdekat antara data dan cluster akan menentukan data mana yang termasuk dalam kelompok data mana. Berikut adalah hasil perhitungan jarak pada iterasi ke-2.

Tabel 9 Hasil Iterasi ke-4

OPD	Jarak ke <i>Centroid</i>			Kelompok
	C1	C2	C3	
Setwan	2.113	5.248	3.665	C1
	80853	50041	34522	
		7		



Inspektorat	3.557	0.643	8.832	C2	Kec. 2X11 Kayutanam	1.205	2.506	6.502	C1
	89144	42680	89734			23600	43265	02790	
Disdikbud	0.874	4.225	4.828	C1	Kec. V Koto Kampung Dalam	2.801	1.103	8.122	C2
	24090	15920	08073			04905	61849	32917	
	9	9	5			9	6	6	
Dinkes	3.979	7.217	1.557	C3	Kec. Ulakan Tapakis	1.638	1.954	6.812	C1
	58316	49765	66350			97828	51953	60417	
DinsoP3A	1.975	2.013	6.885	C1	Kec. Sintoga	4.332	1.046	9.657	C2
	11666	08031	81464			57038	68417	12555	
DLHKP	0.438	3.419	5.429	C1	Kec. Padang Sago	4.392	1.894	9.155	C2
	78145	54517	67214			38835	01736	61759	
DPMD	1.228	4.538	4.266	C1	Batang Gasan	3.850	0.846	9.132	C2
	88926	32572	43374			37433	37414	35491	
DPPKB	0.583	3.110	5.778	C1	Kec. V Koto Timur	2.732	0.730	7.995	C2
	57590	15374	07403			99722	04622	96149	
DISDAGNAK ER	3.460	6.725	2.122	C3	Kec 2x11 Enam Lingkung	5.406	2.034	10.74	C2
	70765	98324	15839			01832	86799	28831	
DPMPTP	1.075	2.577	6.486	C1	Kec Patamuan	3.115	0.938	8.190	C2
	47950	08463	65169			70283	84765	14486	
Dinas Pariwisata	0.439	3.440	5.552	C1	Kec. Enam Lingung	2.956	0.649	8.145	C2
	09053	68042	76269			98593	85563	33677	
Dinas Kearsipan	1.430	3.838	5.988	C1	Kc IV Koto Aur Malintang	5.097	1.832	10.46	C2
	17394	87848	10806			50067	89429	89269	
Dinas Pertanian	1.087	2.853	6.015	C1		8	2	1	
	32968	13042	09880						
Dinas Peternakan	1.462	2.051	6.675	C1		1	1		
	83930	41785	67866						
Dinas PUPR	4.813	8.211	1.986	C3		5	1	1	
	65798	55193	00577						
Dinas Perikanan	2.832	0.874	8.182	C2		2	6	6	
	86642	17768	68775						
Disdukcapil	5.961	9.024	1.597	C3		18675	64366	16840	
	3	6	6						
BPBD	6.761	9.875	1.925	C3		62324	81499	14421	
	3	5	5						
Bappeda- BKPSDM	8.445	11.81	3.142	C3		79348	29801	01998	
	8	5	5						
Kec. Batang Anai	1.052	4.420	4.481	C1		83846	08591	65805	
5	4	4							
Kec. Nan Sabaris	4.943	1.635	10.23	C2		23923	98657	25665	
	8	2	2						

7. Menghitung Titik Pusat Centroid Baru (Iterasi 5)

Pada iterasi pertama data telah menghasilkan keanggotaan cluster, selanjutnya penulis melakukan perhitungan untuk menentukan nilai titik pusat cluster baru yang diperoleh dari menghitung rata-rata keanggotaan dari cluster 1, cluster 2, dan cluster 3, berikut merupakan nilai *centroid* baru:

Tabel 10 Titik Pusat *Cluster* Iterasi ke-5

	Inbound Average	Inbound Maksimum	Outbound Average	Outbound Maksimum
C1	2.726	7.516	0.250	1.404
C2	1.840	4.167	0.240	0.780
C3	5.293	12.115	0.568	3.322

Proses penghitungan nilai *centroid* baru telah dilakukan, langkah selanjutnya adalah menghitung ulang data dengan nilai *centroid* baru. Dalam studi kasus ini, penulis melakukan perhitungan hingga iterasi ke-5. Pada iterasi 4 dan 5 data tidak merubah keanggotaan *cluster*.



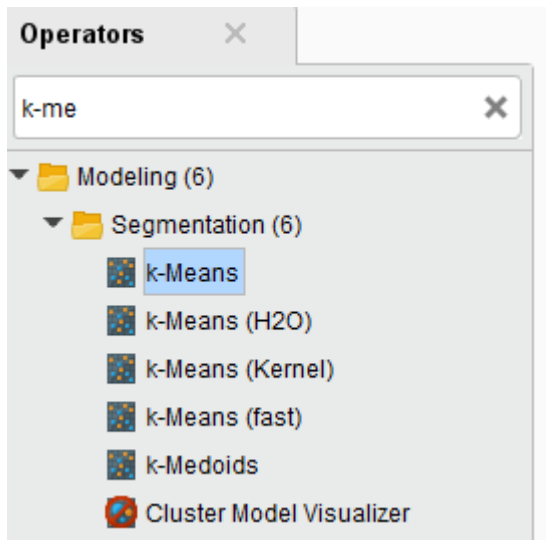
Tabel 11 Hasil Iterasi ke-5

OPD	Jarak ke <i>Centroid</i>			Kelo mpo k
	C1	C2	C3	
Setwan	2.236	5.573	3.665	C1
	50137	31207	34522	
	8	3		
Inspektorat	3.347	0.597	8.832	C2
	76500	88791	89734	
	9	5	9	
Disdikbud	1.078	4.530	4.828	C1
	46457	56237	08073	
	6	8	5	
Dinkes	4.160	7.530	1.557	C3
	80929	27290	66350	
	4	1	1	
DinsoP3A	1.720	2.338	6.885	C1
	08118	48971	81464	
	1			
DLHKP	0.513	3.717	5.429	C1
	23607	35387	67214	
	7		3	
DPMD	1.389	4.855	4.266	C1
	14546	85107	43374	
	7	1	7	
DPPKB	0.544	3.401	5.778	C1
	50028	65140	07403	
	1	7	8	
DISDAGNAKER	3.654	7.033	2.122	C3
	15602	36031	15839	
	7	3	4	
DPMPTP	0.933	2.859	6.486	C1
	73611	90820	65169	
	2	8	8	
Dinas Pariwisata	0.579	3.728	5.552	C1
	13123	82963	76269	
	3	7	8	
Dinas Kearsipan	1.548	4.090	5.988	C1
	78984	39626	10806	
	5	7	9	
Dinas Pertanian	0.976	3.139	6.015	C1
	48150	78695	09880	
	4	5	1	
Dinas Peternakan	1.218	2.368	6.675	C1
	15128	48579	67866	
	6	4	1	
Dinas PUPR	5.024	8.524	1.986	C3
	85605	85253	00577	
	9	8		
Dinas Perikanan	2.630	1.065	8.182	C2
	49229	49438	68775	
	2	8	6	
Disdukcapil	6.127	9.329	1.597	C3
	43931	64136	16840	
	5	3	6	
BPBD	6.912	10.20	1.925	C3
	38356	06537	14421	
	5	2	5	
Bappeda-BKPSDM	8.653	12.12	3.142	C3
	87653	66580	01998	
	8	6	5	
Kec. Batang Anai	1.280	4.720	4.481	C1
	17118	29412	65805	
	8	5	4	
Kec. Nan Sabaris	4.720	1.359	10.23	C2
	78873	89728	25665	
	4	5	2	
Kec. 2X11 Kayutanam	1.003	2.810	6.502	C1
	37650	44607	02790	
	6	2	8	
Kec. V Koto Dalam	2.575	1.349	8.122	C2
	15255	64160	32917	
	8	4	6	
Kec. Ulakan Tapakis	1.384	2.279	6.812	C1
	69898	79475	60417	
	2	5		
Kec. Sintoga	4.119	0.734	9.657	C2
	78026	35955	12555	
	3	7	1	
Kec. Padang Sago	4.135	1.952	9.155	C2
	86795	76461	61759	
	1	7		
Batang Gasan	3.644	0.675	9.132	C2
	47093	71674	35491	
	7	9	3	
Kec. V Koto Timur	2.498	1.047	7.995	C2
	80831	70485	96149	
	3	3	6	
Kec 2x11 Enam Lingkung	5.188	1.714	10.74	C2
	59496	27310	28831	
	4	9	9	
Kec Patamuan	2.891	1.128	8.190	C2
	39718	59970	14486	
	8	4	5	
Kec. Enam Lingung	2.714	0.966	8.145	C2
	06850	40267	33677	
	4	2	9	
Kc IV Koto Aur Malintang	4.889	1.517	10.46	C2
	57941	43642	89269	
	4	1	1	

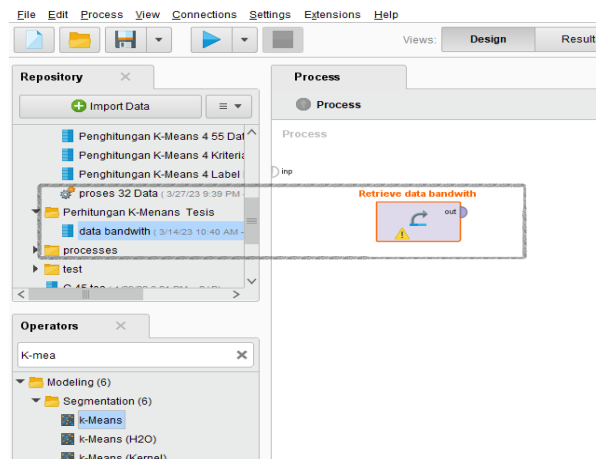
8. Implementasi Algoritma K-Means Menggunakan RapidMiner

Penerapan Algoritma K-Means menggunakan RapidMiner bertujuan untuk mengelompokkan data penggunaan *bandwidth*, untuk memulai pilih operator K-Means



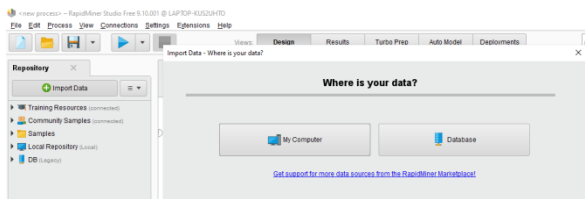


Gambar 2. Memilih Metode K-Means.



Gambar 4. Lembar Proses

Sebelum menerapkan algoritma k-Means maka data dengan format *.xls di import terlebih dahulu ke dalam RapidMiner dengan klik “Add Data”. Selanjutnya pilih lokasi penyimpanan dataset yang akan digunakan. Karena dataset penulis berbentuk file yang memiliki format *.xls dan bukan berasal dari database seperti sql maka pilih lokasi dari “My Computer”.



Gambar 3. Import Data

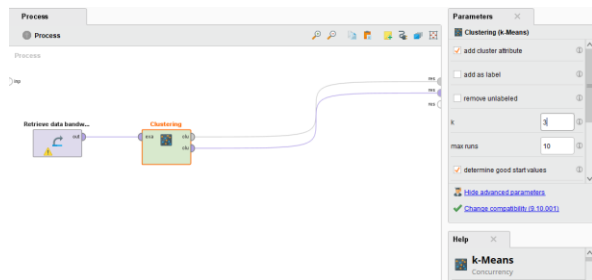
Selanjutnya pilih data import ke dalam *RapidMiner*. Apabila data yang dipilih telah terindikasi tidak ada masalah maka akan muncul “no problems” pada proses import data. Lalu lanjutkan hingga finish, sehingga data berhasil di import ke dalam *RapidMiner*.

9. Membuat Lembar Proses

Setelah data diimpor, data dimasukkan ke halaman Proses. Dengan menyeret data yang diimpor. Contoh file data dalam penelitian yang digunakan disini adalah data bandwidth 2022. Kemudian *drag* dan *drop* bandwidth 2022 ke halaman proses. Gambar 4 di bawah ini merupakan implementasi dari *RapidMiner*.

10. Menerapkan Algoritma K-Means

Langkah selanjutnya adalah memasukkan Algoritma k-Means pada proses yang telah terbentuk diatas untuk mengetahui pengelompokkan dari proses yang telah dibuat. Cari “K-Means” pada Operator lalu tarik ke sebelah “data” lalu hubungkan ke “Clustering” dan “res”. Gambar 5 dibawah ini merupakan implementasi pada Rapid Miner.



Gambar 5. Proses K-Means

Masukan nilai k yang diinginkan, disini peneliti memasukan nilai 3 karena pada kasus ini peneliti mengelompokkannya menjadi 3 cluster. Setelah itu lakukan *running* pada proses untuk mendapatkan hasil *clustering* dari Algoritma K-Means yang diterapkan pada RapidMiner. Hasil *Clustering* dari Algoritma k-Means dapat dilihat pada gambar 6.

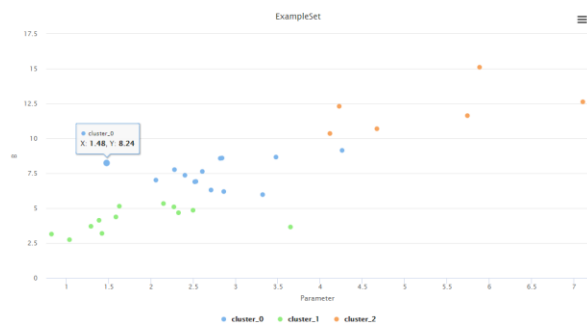


Row No.	id	cluster	Parameter	B	C	D
1	1	cluster_0	4.290	9.140	0.343	1.460
2	2	cluster_1	1.590	4.370	0.178	1.280
3	3	cluster_0	2.820	8.570	0.215	1.200
4	4	cluster_2	4.670	10.690	0.546	3.240
5	5	cluster_0	3.320	5.970	0.281	0.939
6	6	cluster_0	2.610	7.630	0.227	1.890
7	7	cluster_0	3.480	8.660	0.350	1.610
8	8	cluster_0	2.400	7.360	0.222	1.810
9	9	cluster_2	4.120	10.350	0.532	3.220
10	10	cluster_0	2.060	7.010	0.202	0.992
11	11	cluster_0	2.280	7.760	0.325	1.670
12	12	cluster_0	1.480	8.240	0.166	0.844
13	13	cluster_0	2.530	6.910	0.329	2.140
14	14	cluster_0	2.710	6.300	0.236	1.330
15	15	cluster_2	4.230	12.300	0.433	1.660

ExampleSet (32 examples, 2 special attributes, 4 regular attributes)

Gambar 6. Hasil Clustering

Proses *run* akan berjalan dengan baik jika muncul tampilan data pada tampilan *Data View* yang berhasil dijalankan seperti gambar diatas. Pada bagian visualisasi, kita dapat melihat garis dimana ditentukan jumlah *clustering* group di setiap Dinas. Untuk visualisasi di sisi kiri untuk sumbu x, sumbu y, dan warna custom pilih (*cluster*) semuanya akan terlihat pada gambar di bawah ini. Dalam visualisasi tersebut, kita dapat mengetahui bagian grup mana yang mendapatkan posisi *clustering* pada level 0, 1, dan 2.



Gambar 7. Visualisasi

3. Kesimpulan

Berdasarkan penelitian yang telah dilakukan dapat disimpulkan bahwa dengan algoritma K-Means Clustering didapatkan pengetahuan baru (knowledge) dari penggunaan bandwidth di Pemerintah Kabupaten Padang Pariaman sehingga pihak Admin Jaringan di Dinas Komunikasi dan Informatika Kabupaten Padang Pariaman dapat mengoptimalkan pembagian bandwidth berdasarkan tingkat pemakaian bandwidth setiap OPD yang ada di Kabupaten Padang Pariaman. Dari data bandwidth selama tahun 2022, penggunaan bandwidth telah dikelompokkan menjadi 3 kelompok yaitu OPD yang memiliki tingkat penggunaan bandwidth tinggi terdapat 6 OPD, lalu tingkat pemakaian sedang terdapat 14 OPD, sedangkan untuk tingkat pemakaian bandwidth rendah sebanyak 12 OPD.

4. Daftar Pustaka

- [1] J. Konsera, T. Hartati, dan Y.A. Wijaya, "Analisa Perbandingan Manajemen Bandwidth Simple Queue Dan Queue Tree Pada Aplikasi Video Conference Zoom", Jurnal Ilmiah NERO Vol. 7 No. 2, DOI: <http://dx.doi.org/10.21107/nero.v7i2.330>
- [2] Zaenal Mustakim dan Rahmat Kamal, "K-Means Clustering for Classifying The Quality Management of Secondary Education In Indonesia" Cakrawala Pendidikan, Vol. 40, No. 3, October 2021 DOI: <https://doi.org/10.21831/cp.v40i3.40150>
- [3] S. Tongbram, B.A. Shimray, dan L.S Singh "Segmentation of image based on k-means and modified subtractive clustering" Indonesian Journal of Electrical Engineering and Computer Science Vol. 22, No. 3, June 2021, DOI: <http://doi.org/10.11591/ijeecs.v22.i3.pp1396-1403>
- [4] D. Adillah, N. Manurung, dan A. Dermawan, "Implementation Of K-Means Clustering Analysis To Determine Barriers To Online Learning Case Study: Swasta Yependak Tinjowan Junior High School", Jurnal Teknik Informatika (JUTIF) Vol. 3, No. 3, Juni 2022, DOI: <https://doi.org/10.20884/1.jutif.2022.3.3.189>
- [5] F. Valerian dan S. Yulianto, "Identification Of The Covid-19 Distribution Area On The Island Of Kalimantan Using The K-Means Spatial Clustering Method" Jurnal Teknik Informatika (JUTIF) Vol. 3, No. 4, Agustus 2022, hlm. 839-846, DOI: <https://doi.org/10.20884/1.jutif.2022.3.4.314>
- [6] F Nuraeni, D. Tresnawati, Y.H Agustin, dan G.F Dermawan. " Optimization Of Market Basket Analysis Using Centroid-Based Clustering Algorithm And Fp-Growth Algorithm ", Vol. 3, No. 6, Desember 2022, hlm. 1581-1590, Jurnal Teknik Informatika (JUTIF), DOI: <https://doi.org/10.20884/1.jutif.2022.3.6.399>
- [7] D.S. Maylawati Et Al, " Data science for digital culture improvement in higher education using K-means clustering and text analytics ", International Journal of Electrical and Computer Engineering (IJECE) Vol. 10, No. 5, October 2020, DOI: <http://doi.org/10.11591/ijece.v10i5.pp4569-4580>
- [8] Malak Fraihat et al, "An efficient enhanced k-means clustering algorithm for best offer prediction in telecom", International Journal of Electrical and Computer Engineering (IJECE) Vol. 12, No. 3, June 2022, DOI: <http://doi.org/10.11591/ijece.v12i3.pp2931-2943>
- [9] A D Indriyanti, R Prehanto2, dan T Z Vitadiar "K-means method for clustering learning classes", Indonesian Journal of Electrical Engineering and Computer Science Vol. 22, No. 2, May 2021, pp. 835~841, DOI: <http://doi.org/10.11591/ijeecs.v22.i2.pp835-841>
- [10] B.A Jaafar, M.T Gaata, dan M.N Jasim. "Home Appliances Recommendation System Based On



- Weather Information Using Combined Modified K-Means And Elbow Algorithms", Indonesian Journal of Electrical Engineering and Computer Science Vol. 19, No. 3, September 2020, DOI: <http://doi.org/10.11591/ijeecs.v19.i3.pp1635-1642>
- [11] Mourad Fariss *et al*, "A Semantic Web Services Discovery Approach Integrating Multiple Similarity Measures And K-Means Clustering", Indonesian Journal of Electrical Engineering and Computer Science Vol. 24, No. 2, November 2021, DOI: <http://doi.org/10.11591/ijeecs.v24.i2.pp1228-1237>
- [12] Al-Obaydy, Wasseem N. Ibrahim et al, " Document classification using term frequency-inverse document frequency and K-means clustering" Indonesian Journal of Electrical Engineering and Computer Science Vol. 27, No. 3, September 2022, DOI: <http://doi.org/10.11591/ijeecs.v27.i3.pp1517-1524>
- [13] I. Soliani dan S. Juanita, "Grouping The Prevalence Of Disease Cases By Age In Bandung City Hospitals Using K-Means", Jurnal Teknik Informatika (JUTIF), Vol. 3, No. 6, Desember 2022, DOI: <https://doi.org/10.20884/1.jutif.2022.3.6.430>
- [14] N. Puspitasari, J.A. Widians, dan N.B. Setiawan, " Segmentasi pelanggan menggunakan algoritme bisecting k-means berdasarkan model recency, frequency, dan monetary (RFM) " Jurnal Teknologi dan Sistem Komputer, 8(2), 2020, 78-83 DOI: DOI:10.14710/jtsiskom.8.2.2020.78-83
- [15] Y. Sari, A.R Baskara dan P.B Prakoso, "Penerapan Metode K-Means Berbasis Jarak Untuk Deteksi Kendaraan Bergerak", Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK) Vol. 9, No. 6, Agustus 2022, DOI: <http://dx.doi.org/10.21107/nero.v7i2>
- [16] A. Hardianti dan D. Agushinta R, "Analisis Pola Masa Studi Mahasiswa Fakultas Teknik Universitas Darma Persada Menggunakan Metode Clustering K-Means" Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK) DOI: 10.25126/jtiik.202071001 Vol. 7, No. 4, Agustus 2020, DOI: <http://dx.doi.org/10.25126/jtiik.202074>
- [17] Yudi Istianto, "Klasifikasi Kebutuhan Jumlah Produk Makanan Customer Menggunakan K-Means Clustering Dengan Optimasi Pusat Awal Cluster Algoritma Genetika" Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK) DOI: 10.25126/jtiik.202182990 Vol. 8, No. 5, Oktober 2021, DOI: <http://dx.doi.org/10.25126/jtiik.202185>
- [18] S. Widodo, H. Brawijaya, dan Samudi, "Clustering Kanker Serviks Berdasarkan Perbandingan Euclidean dan Manhattan Menggunakan Metode K-Means", Jurnal Media Informatika Budidarma Volume 5, Nomor 2, April 2021, DOI: <http://dx.doi.org/10.30865/mib.v5i2>
- [19] Haviluddin , "Implementasi Metode K-Means untuk Pengelompokan Rekomendasi Tugas Akhir", Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer Vol. 16, No. 1 Februari 2021, DOI: <http://dx.doi.org/10.30872/jim.v16i1>
- [20] S.A. Arnomo dan Y. Yulia "Pengelompokan potensi peningkatan bandwidth pelanggan broadband FTTH" ILKOM Jurnal Ilmiah Vol. 13, No. 1, April 2021, pp. 51-57 <https://doi.org/10.33096/ilkom.v13i1.805.51-57>
- [21] Rubangiya, T. Hartati, dan Y.A Wijaya, "Analisis Data Lalu Lintas Jaringan Di Kantor Cangehgar Cyber Operation Center Menggunakan Algoritma K-Means" Jurnal Ilmiah NERO Vol. 7 No. 1 DOI: 10.21107/nero.v7i1.327

