

PENGELOMPOKAN PENDERITA PENYAKIT COVID-19 MENGGUNAKAN METODE K-MEANS (STUDI KASUS JAWA TIMUR)

Akhmad Fajrul Mujab, Ali Mahmudi, Nurlaily Vendyansyah
Program Studi Teknik Informatika S1, Fakultas Teknologi Industri
Institut Teknologi Nasional Malang, Jalan Raya Karanglo km 2 Malang, Indonesia
1618126scholar.itn.ac.id

ABSTRAK

Informasi penderita covid-19 pada pemerintah provinsi jawa timur dapat dilihat di *website* resmi yaitu <http://infocovid19.jatimprov.go.id/>. Mungkin sebagian dari masyarakat di Indonesia kurang memahami tentang pengelompokan kota mana saja yang terjangkau penyakit *covid-19* karena pada *website* hanya ditampilkan peta dan warna wilayah yang terjangkau saja dan sedikit sulit dipahami oleh masyarakat awam untuk melihat hasil dari setiap daerah yang terjangkau penyakit *covid-19* khususnya di daerah provinsi jawa timur.

Berdasarkan penelitian yang sudah dilakukan maka perlu dibuat suatu sistem yang dibuat menerapkan metode *K-Means Clustering* untuk proses pengelompokan jumlah pasien penderita *covid-19* dan berbasis *website*. Penelitian ini menggunakan data *covid-19* yang diambil dari data *website* pemerintah provinsi jawa timur pada tanggal 10 juni 2020, yang kemudian dilakukan perhitungan dengan metode *k-means*. *Kriteria* yang digunakan pada sistem ini kriteria penderita covid-19 ini dikelompokkan menjadi 4 yaitu orang dalam pantauan (odp), pasien dalam pengawasan (pdp), orang tanpa gejala (otg) dan positif. Pada sistem ini ada dua pengguna yaitu admin dan user, untuk admin bisa melakukan *update* data untuk perhitungan *kmeans* sedangkan *user* hanya bisa melihat hasil perhitungan yang dilakukan oleh *admin*. Penelitian menunjukkan bahwa data pasien covid-19 tanpa perhitungan metode dengan data yang dilakukan perhitungan metode *kmeans* memiliki keakuratan 55%

. Penelitian ini menunjukkan bahwa sistem pengelompokan penderita penyakit covid-19 menggunakan metode *k-means* menghasilkan data *ranking* yang cukup sesuai pengelompokan yang telah ditentukan. Sehingga, sistem ini diharapkan dapat membantu masyarakat dan pemerintah daerah untuk lebih mempersiapkan diri dalam menghadapi pandemi seperti sekarang ini sehingga dapat membantu memudahkan masyarakat pada umumnya untuk mengetahui hasil dari setiap kota yang terjangkau *covid-19* dan lebih waspada dan selalu menjaga kebersihan dan kesehatan dalam kehidupan sehari-hari.

Kata Kunci : *covid-19, kmeans clustering, website, pengelompokan, jawa timur*

1. PENDAHULUAN

Virus Covid-19 adalah bagian dari keluarga *virus* yang menyebabkan penyakit pada hewan ataupun pada manusia. Di Indonesia, masih melawan *virus Corona* hingga saat ini, begitupun juga di negara-negara lain. Jumlah kasus *virus covid-19* terus bertambah dengan beberapa melaporkan kesembuhan, tapi tak sedikit yang meninggal. di Indonesia jumlah kasus *virus corona* yang terkonfirmasi adalah 1792 sembuh 4351 dirawat dan 553 meninggal per tanggal 10 juni pukul 9 malam 2020. [1]

Sistem yang dibuat menerapkan metode *K-Means Clustering* untuk proses pengelompokan jumlah pasien penderita *covid-19* dan berbasis *website*. Sehingga dapat membantu memudahkan pemerintah daerah untuk mengontrol warganya maupun masyarakat pada umumnya untuk lebih waspada dan selalu menjaga kebersihan dan kesehatan. *Kriteria* yang digunakan pada sistem ini kriteria penderita *covid-19* ini dikelompokkan menjadi 5 yaitu orang dalam pantauan (odp), pasien dalam pengawasan (pdp), orang tanpa gejala (otg) dan positif. untuk sangat banyak kasus (resiko tinggi)

dengan kasus 2049+, banyak kasus (resiko sedang) dengan kasus 513-2048, Kasus sedang (resiko rendah) dengan kasus 129-512, 128 sangat sedikit kasus (tidak beresiko) 1-16 sedikit kasus (resiko terkontrol) 17-128.

Dari uraian permasalahan di atas, tercipta sebuah pemikiran untuk mengelompokkan jumlah penderita penyakit *covid-19* dengan metode *k-means*. sistem ini nantinya mampu menyediakan jumlah data-data pasien penderita *covid-19* yang memudahkan pihak pemda maupun masyarakat pada umumnya untuk berhati-hati jika kota yang ditinggalinya banyak terjangkau *covid-19*.

2. TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Beberapa keterkaitan dari penelitian yang dilakukan oleh Bayu Dwiputra Mudzakkir, pada tahun 2018 yang berjudul "pengelompokan data penjualan produk pada pt advanta seeds Indonesia menggunakan metode *k-means*" pesatnya perkembangan teknologi informasi yang menjadikan semua informasi dapat disimpan dalam jaringan computer, membuat munculnya sistem basis data

yang sangat besar. data mining salah satunya, merupakan teknologi yang sangat berguna untuk membantu perusahaan menemukan informasi yang sangat penting dari gudang data (data Warehouse). clustering merupakan salah satu metode dalam data mining yang bersifat tanpa arahan (unsupervised). Penggunaan algoritma k-means adalah dapat membantu dalam mengelompokan data, informasi yang ditampilkan berupa nilai centeroid dari tiap – tiap cluster, untuk mengelompokan barang berdasarkan tingkat kelarisan barang pada pt.advanta seeds indonesia . [2]

Selanjutnya Penelitian dilakukan oleh Tutik Khotimah., pada tahun 2014 yang berjudul “pengelompokan surat dalam al qur’an menggunakan algoritma k-means”. Pada penelitian ini penulis Kitab suci Al Qur’an terdiri dari 114 surat. Masing-masing surat memiliki jumlah ayat yang berbeda. Ayat ayat tersebut ada yang panjang dan ada yang pendek. Banyak atau sedikit jumlah ayat dalam suatu surat, dan panjang atau pendek ayat memberi pengaruh terhadap lama waktu membaca surat-surat dalam Al Qur’an Penelitian ini bertujuan untuk melakukan pengelompokan (clustering) surat-surat dalam Al Qur’an berdasarkan jumlah ayat, jumlah ruku’ dan lama waktu membaca surat. Algoritma clustering yang digunakan adalah K-means. [3]

Kemudian Penelitian yang dilakukan Iin Parlina dkk, tahun 2018 mengenai “ memanfaatkan algoritma k-means dalam menentukan pegawai yang layak mengikuti asesment center untuk clustering program sdp”.Data mining merupakan teknik pengolahan data dalam jumlah besar untuk pengelompokan. Teknik Data mining mempunyai beberapa metode dalam mengelompokkan salah satu teknik yang dipakai penulis saat ini adalah K-Means. Dalam hal ini penulis mengelompokan data daftar program SDP tahun 2017 untuk mengetahui manakah pegawai yang layak lolos dalam program SDP sehingga dapat melakukan Registrasi Asesment Center. Pengelompokan tersebut berdasarkan kriteria – kriteria data Program SDP. [4]

Proses monitoring tumbuh kembang bayi hingga balita di Indonesia menggunakan lembaga posyandu. Monitoring tersebut meliputi pemeriksaan berat badan dan tinggi badan berdasarkan umur serta pemberian imunisasi secara teratur. Pemeriksaan pertumbuhan batita di lakukan setiap bulan oleh petugas posyandu secara manual. Dalam menentukan status gizi batita, petugas masih kesulitan dalam merekap semua data dengan banyaknya data kasus batita yang semakin bertambah.

Penelitian ini bertujuan untuk membantu petugas dalam mereka dan menentukan status gizi batita. Metode ini meyelesaikan suatu masalah dengan membandingkan permasalahan yang pernah terjadi sebelumnya.Hasil pengujian menunjukkan bahwa beberapa fungsi yang terdapat dalam aplikasi dapat berjalan dengan baik. Dan hasil perbandingan menunjukkan terdapat perbedaan nilai pada masing-

masing variabel perbandingan yang dilakukan pada ketiga metode tersebut, diantaranya jumlah cluster, nilai normalisasi dan jumlah iterasi yang terjadi selama proses perhitungan metode[5].

2.2 Data mining

Merupakan salah satu solusi untuk menjelaskan proses penggalian informasi dalam suatu basis data yang berskala besar dan proses klasifikasi otomatisasi kasus berdasarkan pola data yang diperoleh dari data set. Data mining dapat diterapkan pada berbagai bidang, dibagi menjadi empat kelompok : model prediksi (prediction modelling), analisis cluster (cluster analysis), analisis asosiasi (association analysis), dan deteksi anomaly (anomaly detection). [6]

2.3 K-Means Clustering

K-means clustering merupakan salah satu metode data clustering non-hirarki yang mengelompokan data dalam bentuk satu atau lebih cluster atau kelompok. Data-data yang memiliki karakteristik yang sama dikelompokan dalam satu cluster atau kelompok dan data yang memiliki karakteristik yang ber berbeda dikelompokan dengan cluster atau kelompok yang lain sehingga data yang berada dalamsatu cluster/kelompok memiliki tingkat variasi yang kecil.

Langkah-langkah melakukan clustering dengan metode K-Means adalah sebagai berikut:

- a. Pilih jumlah cluster k.
- b. Inisialisasi k-pusat cluster ini bisa dilakukan dengan berbagai cara. Namun yang paling sering dilakukan adalah dengan cara random. Pusat pusat cluster diberi nilai awal dengan angka-angka random.
- c. Alokasikan semua data/ objek ke cluster terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke cluster tertentu ditentukan jarak antara data dengan pusat cluster. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat cluster. Jarak paling antara satu data dengan satu cluster tertentu akan menentukan suatu data masuk dalam cluster mana. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan seperti ditunjukkan pada persamaan 2.1

$$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2} \dots \dots \dots (2.1)$$

Dimana :

- D(i, j) = Dimana 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

- d. Hitung kembali pusat *cluster* dengan keanggotaan *cluster* yang sekarang. Pusat *cluster* adalah rata-rata dari data/objek dalam *cluster* tertentu. Jika dikehendaki bisa juga menggunakan *median* dari *cluster* tersebut. Jadi rata-rata (*mean*) bukan satu satunya ukuran yang bisa dipakai.
- e. Hitung kembali jarak setiap data dengan pusat *cluster* yang baru seperti langkah no c sampai anggota *cluster* tidak berubah lagi.[7]

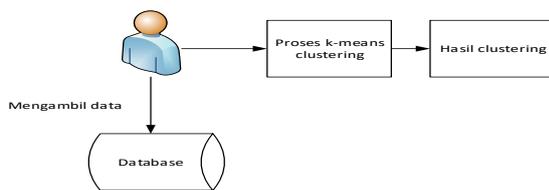
3. METODE PENELITIAN

3.1 Perancangan sistem

Perancangan konsep berguna untuk memberikan solusi terbentuknya pengelompokan yang di analisa dan di proses dengan *Algoritma K-Means Clustering* menggunakan data pasien *covid-19* yang diambil dari website pemrov jatim yaitu <http://infocovid19.jatimprov.go.id/>.

3.2 Blok Diagram Sistem

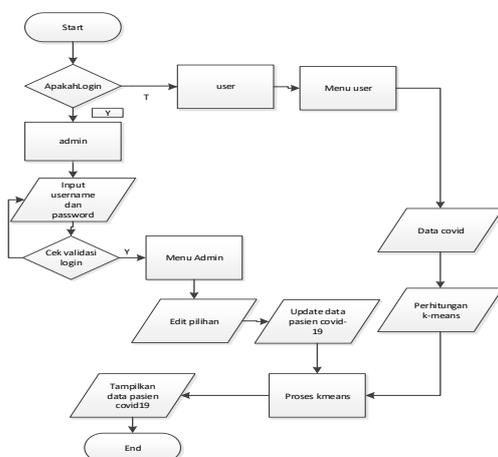
Blok diagram adalah *diagram* dari sebuah sistem, di mana bagian utama atau fungsi yang diwakili oleh blok dihubungkan dengan garis, yang menunjukkan hubungan dari blok. proses kerja pada pengelompokan penderita *covid-19* menggunakan metode *k-means* ditunjukkan pada Gambar 1.



Gambar 1 Blok diagram sistem

3.3 Flowchart Sistem

Flowchart sistem ini menjelaskan proses berjalannya aplikasi seperti ditunjukkan pada Gambar 2

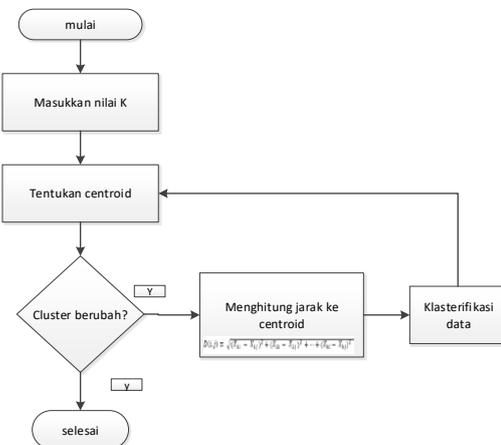


Gambar 2 Flowchart Sistem

Pada gambar 2 *flowchat* sistem menjelaskan tentang alur dari sistem dimulai dari *login* jika login maka otomatis akan langsung masuk pada beranda admin. untuk admin proses *login* akan dicek dahulu apakah benar *password* dan *username*nya sudah benar, kalau sudah benar maka *admin* bisa untuk melakukan pengeditan pada pasien penderita *covid-19*, kemudian bisa untuk melakukan proses *clustering* menggunakan metode *kmeans* dan bisa untuk menampilkan data agar bisa dilihat oleh user biasa. jika tidak *login* maka akan masuk pada user, pada menu *user* ini hanya bisa melihat tampilan data pasien *covid-19* dan melihat tampilan hasil dari data perhitungan *kmeans* yang telah dilakukan oleh *admin*.

3.4 Flowchat Metode

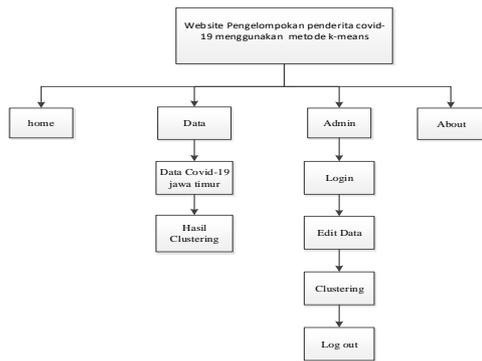
Pada Gambar 3 *Flowchart* program menjelaskan tentang alur dari metode *K-Means Clustering*, di mana metode ini harus menentukan banyaknya *cluster* dulu, setelah jumlah *cluster* di tentukan maka akan di lanjutkan dengan menentukan titik tengahnya, setelah menemukan titik tengahnya maka proses *K-Means Clustering* selesai, jika titik tengahnya berubah maka proses akan di lanjutkan dengan cara mencari jarak terdekat dan mengelompokkan data berdasarkan jarak terdekat, apabila jarak terdekatnya sudah ditemukan maka proses *K-Means clustering* selesai seperti ditunjukkan pada Gambar 3.



Gambar 3 Blok diagram sistem

3.5 Struktur Menu

Struktur menu dirancang untuk menyederhanakan suatu sistem yang telah dibangun untuk melihat semua arsitektur sistem yang telah dirancang, tampilan struktur menu yang telah dirancang dapat dilihat pada Gambar 4



Gambar 4 struktur menu

Penjelasan lebih dari Gambar adalah sebagai berikut:

1. menu home : yaitu tampilan home pada website
2. data :
 - a. data penderita covid-19 di jawa timur diambil dari website pemerintah provinsi jawa timur
 - b. hasil clustering : untuk menampilkan data setelah di clustering
3. admin : merupakan fitur admin yang bisa mengedit data dan melakukan clustering
 - a. login : halaman khusus untuk admin
 - b. edit data : edit data yang dilakukan oleh admin dari data penderita covid-19
 - c. clustering : proses pengelompokan data yang dilakukan admin
 - d. logout : untuk kembali ke menu user
4. about : merupakan profil dari pembuat website

3.6 Data training

Berdasarkan data pasien covid-19, terdapat beberapa kriteria yaitu orang tanpa gejala (otg) (K1), orang dalam pantauan (odp) (K2), pasien dalam pengawasan (pdp) (K3), positif (K4), penderita covid-19 pertanggal 10 juni 2020. Berikut adalah data jumlah pasien. Ditunjukkan pada Tabel 1.

Tabel 1 Data jumlah pasien covid-29 jawa timur

No.	Kabupaten/Kota	otg	odp	pdp	positif	sembuh	mninggal
1	Kab. Bangkalan	833	966	32	117	217	44
2	Kota Surabaya	4007	4040	3512	3627	4733	743
3	Kab. Pasuruan	228	278	183	118	340	64
4	Kab. Jombang	615	660	55	103	322	40
5	Kab. Sidoarjo	1153	1193	553	818	1673	175
6	Kab. Situbondo	527	444	47	75	167	11
7	Kab. Pacitan	567	626	18	16	30	2
8	Kab. Lamongan	787	534	190	154	213	44
9	Kab. Gresik	280	1220	339	247	843	140
10	Kab. Sampang	85	506	20	59	143	12
11	Kab. Malang	302	495	343	110	239	42
12	Kab. Mojokerto	93	568	100	69	251	20
13	Kab.	1198	478	64	120	153	6

	Probolinggo						
14	Kab. Bojonegoro	363	255	24	69	168	24
15	Kota Probolinggo	471	343	23	24	153	6
16	Kab. Tuban	628	639	77	57	123	16
17	Kota Malang	498	933	277	73	172	64
18	Kab. Ponorogo	275	448	54	34	122	5
19	Kab. Lumajang	497	400	111	49	55	11
20	Kab. Trenggalek	192	890	35	14	54	0
21	Kota Mojokerto	52	518	13	21	163	12
22	Kota Kediri	475	334	36	55	167	19
23	Kab. Nganjuk	1069	97	98	38	102	24
24	Kab. Jember	1040	1285	192	66	140	8
25	Kota Blitar	104	209	6	4	24	2
26	Kab. Tulung Agung	805	1330	649	66	239	3
27	Kab. Banyuwangi	157	870	40	11	39	2
28	Kab. Blitar	247	955	54	10	60	11
29	Kab. Ngawi	379	335	41	16	26	0
30	Kab. Madiun	929	344	60	33	14	0
31	Kab. Pamekasan	820	511	73	52	149	27
32	Kab. Kediri	1784	1032	198	172	55	3
33	Kab. Sumenep	35	346	11	16	160	8
34	Kab. Bondowoso	138	1087	10	19	34	1
35	Kab. Magetan	174	337	82	87	108	5
36	Kota Madiun	128	71	19	6	14	0
37	Kota Pasuruan	120	115	13	33	94	20
38	Kota Batu	287	303	83	39	133	9

4. HASIL DAN PEMBAHASAN

4.1. Tampilan home



Gambar 5 tampilan home

Pada Gambar 5 tampilan home merupakan tampilan awal pada website penderita covid-19 di jawa timur, pada tampilan home terdapat menu data, login dan about.

4.2. Tampilan Login



Gambar 7 tampilan login

Pada Gambar 7 tampilan login digunakan untuk masuk pada halaman admin yang nantinya bisa melakukan update tampilan data pada website.

4.3. Halaman Admin



Gambar 8 halaman admin

Pada Gambar 8 tampilan awal pada halaman admin yang nantinya akan bisa mengedit perhitungan *k-means* dan jumlah data pasien covid-19

4.4. Tampilan data admin

Kabupaten/Kota	Orang Tanpa Gejala (OTG)	Orang Dalam Pengawasan (ODP)	Pasien Dalam Pengawasan (PDP)	Positif	Sembuh	Meninggal	Action
Kabupaten Negeri	1009	97	90	38	102	24	EDIT
Kabupaten Bangkalan	966	32	833	117	217	44	EDIT
Kabupaten Banyuwangi	970	40	157	11	39	2	EDIT
Kabupaten Blitar	955	54	247	10	60	11	EDIT
Kabupaten Bojonegara	235	24	363	69	168	24	EDIT
Kabupaten Bondowoso	1087	10	138	19	34	1	EDIT
Kabupaten Gresik	1220	339	280	247	843	140	EDIT
Kabupaten Jember	1285	192	174	87	188	5	EDIT
Kabupaten Jombang	660	35	615	103	322	40	EDIT
Kabupaten Kediri	1034	198	1784	172	197	19	EDIT
Kabupaten Lamongan	534	190	787	154	213	44	EDIT
Kabupaten Lumajang	400	111	497	49	55	11	EDIT
Kabupaten Madiun	344	80	929	33	35	1	EDIT
Kabupaten Magetan	237	174	174	87	188	5	EDIT
Kabupaten Malang	495	343	302	110	239	42	EDIT
Kabupaten Mojokerto	568	100	93	69	251	20	EDIT
Kabupaten Ngawi	335	41	379	16	26	0	EDIT
Kabupaten Pacitan	626	18	567	16	38	2	EDIT
Kabupaten Ponorogo	511	73	820	52	149	27	EDIT

Gambar 9 tampilan data admin

Pada Gambar 9 tampilan data pada halaman admin yang isinya yaitu jumlah data penderita covid-19 dan bisa diedit oleh khusus admin

4.5. Menentukan Jumlah Cluster dan Centroid

Tahap awal perhitungan data menggunakan metode *K-Means* adalah dengan menentukan jumlah k cluster yang akan dibentuk. Pada perhitungan awal ini akan dilakukan perhitungan dengan nilai k sebanyak 5, kemudian diperlukan penentuan centroid awal terlebih dahulu.

Tabel 2 adalah centroid awal.

Centroid	K1	K2	K3	K4
C1	120	115	13	33
C2	128	71	19	6
C3	5671	626	18	16
C4	1153	1193	553	818
C5	4007	4040	3512	3627

Proses penentuan *centroid baru* diambil secara acak pada tabel data pasien covid 19 c1 diambil dari data otg,odp,pdp dan positif dari kota pasuruan,c2 diambil dari data kota madiun,c3 diambil dari data kabupaten pacitan,c4 diambil dari data o kab.sidoarjo dan c5 diambil dari data kota surabaya.

Kemudian menghitung jarak setiap data ke centroid proses ini bertujuan untuk mencari jarak terpendek antara objek ke *centroid* dengan persamaan perhitungan *Euclidean Distance(1)*

$$d(x_1, c_1) = \sqrt{(833 - 120)^2 + (966 - 115)^2 + (32 - 13)^2 + (117 - 33)^2} = 1113.55$$

$$d(x_1, c_2) = \sqrt{(833 - 128)^2 + (966 - 71)^2 + (32 - 19)^2 + (117 - 6)^2} = 1144.79$$

$$d(x_1, c_3) = \sqrt{(833 - 567)^2 + (966 - 626)^2 + (32 - 18)^2 + (117 - 16)^2} = 443.57$$

$$d(x_1, c_4) = \sqrt{(833 - 1153)^2 + (966 - 1193)^2 + (32 - 553)^2 + (117 - 818)^2} = 957.48$$

$$d(x_1, c_5) = \sqrt{(833 - 4007)^2 + (966 - 4040)^2 + (32 - 3512)^2 + (117 - 3627)^2}$$

Hasil Perhitungan pada *centroid* tersebut dilakukan sebanyak data pada tabel data jumlah penderita covid-19 dengan menggunakan 5 *centroid*. Selanjutnya untuk Data ke-2 sampai Data ke-38 dilakukan perhitungan sesuai dengan Data ke-1. Hasil perhitung jarak semua data *centroid* pada Tabel 3

Tabel 3 mencari jarak terpendek antara objek ke centroid

Data Ke-	C1	C2	C3	C4	C5	Cluster
1	1113.55	1144.79	443.57	957.48	6629.80	3
2	7461.52	7490.81	6981.14	5735.58	0.00	5
3	272.69	303.79	523.12	1523.07	7199.22	1
4	740.75	771.23	111.35	1154.45	6877.46	3
5	1771.16	1803.89	1262.66	0.00	5735.58	4
6	526.13	551.25	197.60	1327.01	7047.31	3
7	679.15	707.71	0.00	1262.66	6981.14	3
8	816.35	836.54	324.80	1068.13	6764.45	3
9	1182.67	1226.29	769.16	1065.22	6582.95	3
10	393.49	440.32	498.57	1572.50	7266.08	1
11	540.70	570.83	449.27	1325.43	6978.96	3
12	463.47	508.69	487.41	1510.11	7188.55	1
13	1141.94	1151.33	658.03	1113.36	6690.75	3
14	282.96	305.08	426.73	1531.27	7240.89	1
15	418.77	438.15	298.99	1448.78	7164.21	3
16	733.02	760.65	95.14	1178.24	6894.10	3
17	939.84	975.19	411.51	1061.99	6712.87	3
18	369.59	407.12	344.34	1479.72	7190.11	3
19	482.92	504.69	256.35	1358.64	7066.72	3
20	778.88	821.69	458.93	1389.27	7044.15	3
21	408.87	453.70	526.25	1610.79	7300.19	1
22	418.33	438.48	309.15	1430.74	7145.01	3
23	952.98	945.21	733.98	1422.57	6979.64	3
24	1499.48	1529.40	831.14	846.79	6332.16	3
25	99.91	140.69	623.33	1740.82	7438.29	1
26	1533.31	1563.30	976.18	845.34	6202.77	4
27	756.71	799.82	477.65	1418.01	7071.18	3
28	850.85	892.67	460.41	1333.92	6978.20	3
29	341.40	365.08	347.21	1496.86	7209.48	1
30	842.10	847.67	461.11	1276.81	6925.87	3
31	806.71	823.10	285.58	1180.33	6871.55	3
32	1913.98	1930.14	1304.86	983.58	6075.31	4
33	246.74	290.58	601.23	1704.20	7397.21	1
34	972.27	1016.17	629.79	1405.24	6997.93	3
35	244.70	288.79	497.10	1564.41	7259.52	1
36	52.58	0.00	707.71	1803.89	7490.81	2
37	0.00	52.58	679.15	1771.16	7461.52	1
38	261.09	290.33	432.99	1539.41	7241.23	1

Hasil perhitungan pada *centroid* tersebut dilakukan sebanyak 38 data data penderita covid-19 dan sebanyak 5 cluster. Dapat dilihat pada tabel 4.

Tabel 4 hasil rata rata dari perhitungan cluster satu literasi pertama

Hitung Data	1920	3770	576	531
Jumlah Data	11	11	11	11
Rata-rata	174.5	342.7	52.4	48.3

setelah hasil diperoleh dari iterasi pertama, langkah berikutnya adalah menghitung jumlah rata-rata data sesuai dengan hasil *cluster* yang digunakan untuk menentukan nilai *centroid* terbaru, berikut hasil dari perhitungan dari cluster 1. Di bawah ini merupakan rumus untuk mencari nilai rata-rata yang didapat dari literasi pertama cluster 1 :
=6629.80

$$\frac{\text{hitung data}}{\text{jumlah data}}$$

Dengan hasil berikut :

$$C1K1 = \frac{1920}{11} = 174.5 \quad C1K2 = \frac{3770}{11} = 342.7 \quad C1K3 = \frac{576}{11} = 52.4 \quad C1K4 = \frac{531}{11} = 48.3$$

$$C2K1 = \frac{128}{1} = 128 \quad C2K2 = \frac{71}{1} = 71 \quad C2K3 = \frac{19}{1} = 19 \quad C2K4 = \frac{6}{1} = 6$$

$$C3K1 = \frac{12545}{22} = 570.2 \quad C3K2 = \frac{14559}{22} = 661.8 \quad C3K3 = \frac{2228}{22} = 101.3 \quad C3K4 = \frac{1477}{2} = 67.1$$

$$C4K1 = \frac{3742}{1} = 3742 \quad C4K2 = \frac{3555}{1} = 3555 \quad C4K3 = \frac{1400}{1} = 1400 \quad C4K4 = \frac{1056}{1} = 1056$$

$$C5K1 = \frac{4007}{1} = 4007 \quad C5K2 = \frac{4040}{1} = 4040 \quad C5K3 = \frac{3512}{1} = 3512 \quad C5K4 = \frac{3627}{1} = 3627$$

Hasil pembetulan *centroid* baru pada literasi selanjutnya yang diambil dari hasil rata-rata sesuai *cluster* diliterasi sebelumnya berikut hasil *centroid* baru pada Tabel 5

Tabel 5 centroid baru

Centroid	K1	K2	K3	K4
C1	174.5	342.7	52.4	48.3
C2	178	21	19	6
C3	570.2	661.8	101.3	67.1
C4	1247.3	1185	466.7	352
C5	4007	4040	3512	3627

Apabila hasil centroid yang terbaru berubah, maka perlu di lakukan iterasi lagi sampai *centroid* tidak ada yang berubah daripada *centroid* sebelumnya.

4.6. cluster 1

hasil perhitungan k-means terdapat pada tabel 5.

Tabel 5. Hasil cluster 1

No.	Kabupaten/Kota	Odp	Pdp	Otg	Positip
1	Kab. Pasuruan	228	278	183	118
2	Kab. situbondo	527	444	47	75
3	Kab. Sampang	85	506	20	59
4	Kab. Malang	302	495	343	110
5	kab. Mojokerto	93	568	100	69
6	Kab Bojonegoro	363	255	24	69
7	Kota Probolinggo	471	343	23	24
8	Kab. Ponorogo	275	448	54	34
9	Kab. Lumajang	497	400	111	49
10	Kota Mojokerto	52	518	13	21
11	Kota Kediri	475	334	36	55
12	Kab. Ngawi	379	335	41	16
13	Kab. Magetan	174	337	82	87
14	Kota Batu	287	303	83	39

Hasil dari cluster 1 adalah cluster dengan sedikit kasus(resiko terkontrol) dari pasien covid-1

4.7. cluster 2

Hasil perhitungan k-means terdapat pada tabel 6

Tabel 6. Hasil cluster 2

No.	Kabupaten/Kota	Odp	Pdp	Otg	Positip
1	Kota Madiun	128	71	19	6
2	Kota Pasuruan	120	115	13	33
3	Kota Blitar	104	209	6	4
4	Kab. Sumenep	35	346	11	16

Hasil dari cluster 1 adalah cluster dengan sangat sedikit dari pasien covid-19

4.8. cluster 3

Hasil perhitungan k-means terdapat pada tabel

7.

Tabel 7. Hasil cluster 3

No.	Kabupaten/Kota	Odp	Pdp	Otg	positip
1	Kab. Madiun	929	344	60	33
2	Kab. Pamekasan	820	511	73	52
3	Kab. Banyuwangi	157	870	40	11
4	Kab. Blitar	247	955	54	10
5	Kab. Nganjuk	1069	97	98	38
6	Kab. Trenggalek	192	890	35	14
7	Kab. Tuban	628	639	77	57
8	Kota Malang	498	933	277	73
9	Kab. Probolinggo	1198	478	64	120
10	Kab. Bondowoso	138	1087	10	19
11	Kab. Pacitan	567	626	18	16
12	Kab. Lamongan	787	534	190	154
13	Kab. Gresik	280	1220	339	247
14	Kab. Jombang	615	660	55	103
15	Kab. Bangkalan	833	966	32	117

Hasil dari cluster 3 adalah *cluster* dengan kasus sedang dari pasien covid-19

4.9. cluster 4

Hasil perhitungan k-means terdapat pada tabel

8.

Tabel 8. Hasil cluster 4

No.	Kabupaten/Kota	Odp	Pdp	Otg	positip
1	Kab. Kediri	1784	1032	198	172
2	Kab. Tulungagung	805	1330	649	66
3	Kab. Jember	1040	1285	192	66
4	Kab. Sidoarjo	1153	1193	553	818

Hasil dari *cluster* 4 adalah cluster dengan banyak kasus dari pasien covid-19

4.10. cluster 5

Hasil perhitungan k-means terdapat pada tabel 9.

Tabel 9. Hasil cluster

No.	Kabupaten/Kota	Odp	Pdp	Otg	positip
1	Kota Surabaya	4007	4040	3512	3627

Hasil dari *cluster* 5 adalah cluster dengan sangat banyak dari pasien covid-19

4.11. Pengujian fungsional

fungsional sistem dilakukan untuk menguji fitur-fitur yang ada pada sistem pengelompokan jumlah penderita covid-19. Hasil

pengujian sistem ditampilkan pada tabel 10, pengujian fungsional menunjukkan hasil pengujian sistem yang telah dilakukan mendapatkan hasil bahwa sebagaimana fungsi pada sistem telah berjalan dengan baik.

Tabel 10 pengujian fungsional

NO	Modul Fungsi	Mozilla Firefox versi 78.0.2	Internet Explore versi 11.0
1.	Masuk ke system dengan akses sebagai Admin	Berhasil	Berhasil
	1. Admin dapat mengakses menu home	Berhasil	Berhasil
	2. Login dengan memasukkan username dan pasword sesuai hak akses	Berhasil	Berhasil
	3. Admin dapat mengupdate data penderita covid-19	Berhasil	Berhasil
	4. Admin dapat mengupdate data covid-19	Berhasil	Berhasil
	5. Admin dapat melakukan perhitungan K-Means	Berhasil	Berhasil
	6. Admin dapat mengupdate centroid	Berhasil	Berhasil
	7. Admin dapat melakukan perhitungan iterasi	Berhasil	Berhasil
	8. Admin dapat mengakses halaman about	Berhasil	Berhasil
2.	Masuk ke sistem sebagai user	Berhasil	Berhasil
	1. User dapat mengakses menu home	Berhasil	Berhasil
	2. User dapat melihat data covid	Berhasil	Berhasil
	3. User dapat melihat proses perhitungan k-means	Berhasil	Berhasil
	4. User dapat melihat hasil perhitungan kmeans	Berhasil	Berhasil
	5. User dapat mengakss menu about	Berhasil	Berhasil

Tabel 11 perbandingan hasil

No	Kabupaten /kota	Hasil		Keterangan
		Data pada website http://infocovid19.jatimprov.go.id/	Sistem k-means	
1	Kabupaten Nganjuk	sedikit kasus	kasus sedang	Hasil Sama
2	Kabupaten Bangkalan	sedikit kasus	kasus sedang	Hasil Sama
3	Kabupaten Banyuwangi	sangat sedikit kasus	kasus sedang	Hasil Sama
4	Kabupaten Blitar	sangat sedikit kasus	kasus sedang	Hasil Sama
5	Kabupaten Bojonegoro	sedikit kasus	sedikit kasus	Hasil Sama
6	Kabupaten Bondowoso	sedikit kasus	kasus sedang	Hasil Sama
7	Kabupaten Gresik	kasus sedang	kasus sedang	Hasil Sama
8	Kabupaten Jember	sedikit Kasus	banyak kasus	Hasil Sama
9	Kabupaten Jombang	sedikit kasus	kasus sedang	Hasil Sama

10	Kabupaten Kediri	kasus sedang	banyak Kasus	Hasil Sama
11	Kabupaten Lamongan	kasus sedang	kasus sedang	Hasil Sama
12	Kabupaten Lumajang	sedikit kasus	sedikit Kasus	Hasil Sama
13	Kabupaten Madiun	sedikit kasus	kasus sedang	Hasil Sama
14	Kabupaten Magetan	sedikit kasus	sedikit Kasus	Hasil Sama
15	Kabupaten Malang	sedikit kasus	sedikit Kasus	Hasil Sama
16	Kabupaten Mojokerto	sedikit kasus	sedikit Kasus	Hasil Sama
17	Kabupaten Ngawi	sedikit kasus	sedikit Kasus	Hasil Sama
18	Kabupaten Pacitan	sangat sedikit kasus	kasus sedang	Hasil Sama
19	Kabupaten Pamekasan	sedikit kasus	kasus sedang	Hasil Sama
20	Kabupaten Pasuruan	sedikit kasus	sedikit Kasus	Hasil Sama
21	Kabupaten Ponorogo	sedikit kasus	sedikit kasus	Hasil Sama
22	Kabupaten Probolinggo	sedikit kasus	kasus sedang	Hasil Sama
23	Kabupaten Sampang	sedikit kasus	sedikit kasus	Hasil Sama
24	Kabupaten Sidoarjo	banyak Kasus	banyak Kasus	Hasil Sama
25	Kabupaten Sumenep	sangat sedikit kasus	sangat sedikit kasus	Hasil Sama
26	Kabupaten Trenggalek	sangat sedikit kasus	kasus sedang	Hasil Sama
27	Kabupaten Tuban	sedikit kasus	kasus sedang	Hasil Sama
28	Kabupaten Tulung agung	sedikit kasus	banyak kasus	Hasil Sama
29	Kota Batu	sedikit kasus	sedikit kasus	Hasil Sama
30	Kota Kediri	sedikit kasus	sedikit kasus	Hasil Sama
31	Kota Madiun	sangat sedikit kasus	sangat sedikit kasus	Hasil Sama
32	Kota Malang	sedikit kasus	kasus sedang	Hasil Sama
33	Kota Mojokerto	sedikit kasus	kedikit kasus	Hasil Sama
34	Kota Pasuruan	sedikit kasus	sangat sedikit kasus	Hasil Sama
35	Kota Probolinggo	sedikit kasus	sedikit kasus	Hasil Sama
36	Kota Surabaya	sangat banyak kasus	sangat sedikit kasus	Hasil Sama
37	Kabupaten situbondo	sedikit kasus	sedikit kasus	Hasil Sama
38	Kota blitar	sangat sedikit kasus	sangat sedikit kasus	Hasil Sama

Berdasarkan Tabel 11 dapat dilihat persentase untuk *Correctly Clustering* adalah sebesar 55% jumlah data tidak sama dibagi jumlah data kemudian dikalikan 100 jumlah data yang tidak sama dibagi jumlah data covid-19 dengan hasil perhitungan sebagai berikut :

$$\frac{21 \text{ data yang tidak sama}}{38 \text{ data covid-19}} \times 100 = 55\%$$

Dimana dari 38 data kota kabupaten yang terjangkau covid, ada sebanyak 20 data berhasil di

clustering dengan benar dan sebanyak 18 data tidak berhasil di clustering dengan benar sangat banyak kasus (resiko tinggi) dengan kasus 2049+, banyak kasus (resiko sedang) dengan kasus 513-2048, Kasus sedang (resiko rendah) dengan kasus 129-512 sangat sedikit kasus(tidak beresiko)1-16 sedikit kasus (resiko terkontrol)17- 128.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan perhitungan dengan metode k-means dengan langkah-langkah sebagai berikut :

1. Hasil dari perhitungan metode k-means clustering yang dibagi menjadi 5 cluster. sangat banyak kasus (zona hitam) dengan kasus terkonfirmasi positif 2049+, banyak kasus (zona merah tua) dengan kasus terkonfirmasi positif 513-2048, Kasus sedang (zona merah bata) dengan kasus terkonfirmasi positif 129-512 sangat sedikit kasus (zona merah muda) terkonfirmasi positif 1-16 sedikit kasus (zona merah) terkonfirmasi positif 17- 128. Jumlah data pada cluster 1 sebanyak 14 data, cluster 2 sebanyak 4 data, cluster 3 sebanyak 14 data dan cluster 5 sebanyak 1 data.
2. Berdasarkan perbandingan hasil dapat dilihat persentase untuk Correctly Clustering adalah sebesar Berdasarkan perbandingan hasil dapat dilihat persentase untuk Correctly Clustering adalah sebesar 45% sementara persentase untuk Incorrectly Clustering adalah sebesar 55%. Dimana dari 38 jumlah data penderita covid-19, ada sebanyak 22 data berhasil di clustering dengan benar dan sebanyak 16 data tidak berhasil di clustering dengan benar, oleh karena itu untuk mendapatkan keakuratan data dengan data website pemrov jatim perlu penambahan metode lain.
3. Berdasarkan hasil pengujian fungsional melalui 2 browser, Sistem Pengelompokan jumlah pasien covid-19, 100% fungsi sudah berjalan sesuai dengan apa yang telah diharapkan.

5.2 Saran

Adapun saran-saran yang diperlukan untuk menindaklanjuti penelitian dan perbaikan untuk kedepannya adalah sebagai berikut :

1. Sitem dapat membaca data pasien covid19 secara update sesuai data yang ada pada website pemerintah provinsi jatim
2. Pengembangan selanjutnya, menambahkan Grafik.
3. Penelitian selanjutnya dilakukan dengan perbandingan metode K-means dan (KNN) untuk mengetahui nilai keakuratan yang maksimal dengan menggunakan data yang sama.

DAFTAR PUSTAKA

- [1] pemrovjatim. (2020, April 10). Retrieved from infocovid19: <http://infocovid19.jatimprov.go.id/>
- [2] bayu dwiputra, m. (2018). "Pengelompokan Data Penjualan Produk Pada Pt Advanta Seeds Indonesia Menggunakan Metode K - Means" *JATI (Jurnal Mahasiswa Teknik Informatika)*, 2(2)vol 2 no 2
- [3] Khotimah, T. (2014). Pengelompokan Surat Dalam Al Qur'an Menggunakan Algoritma K-Means. *Jurnal Simetris*, 5(1), 83-88..
- [4] Parlina, I., Windarto, A. P., Wanto, A., & Lubis, M. R. (2018). Memanfaatkan Algoritma K-Means dalam Menentukan Pegawai yang Layak Mengikuti Asessment Center untuk Clustering Program SDP. *Computer Engineering, Science and System Journal*, 3(1), 87-93..
- [5] Effendi, M. (2018). Penerapan Metode K-Means Clustering Pada Sistem Pendukung Keputusan Status Gizi Batita, 184-189.
- [6] Darmi, Y. D., & Setiawan, A. (2016). Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk. *Jurnal Media Infotama*, 12(2).
- [7] Abdillah, G., Putra, F. A., Renaldi, F., & Informatika, P. S. (2016). Penerapan Data Mining Pemakaian Air Pelanggan Untuk Menentukan Klasifikasi Potensi Pemakaian Air Pelanggan Baru Di Pdam Tirta Raharja Menggunakan Algoritma K-Means. *vol. 1, SENTIKA*, 498-506.