

**ANALISA PENGARUH JUMLAH TENAGA MEDIS TERHADAP
PENYEBARAN KASUS COVID-19
PADA KABUPATEN MALANG
MENGUNAKAN METODE K-MEANS CLUSTERING**

Bagus Java

Magister Informatika, Fakultas Sains dan Teknologi, UIN Maulana Malik Ibrahim, Malang 65145, Indonesia

Email : bagusjava7@gmail.com

ABSTRAK

Penyakit corona virus 2019 atau Corona Virus Disease-19 (COVID-19) adalah infeksi saluran pernapasan yang disebabkan oleh jenis virus corona. Kasus COVID-19 pertama kali dilaporkan di Kota Wuhan, Provinsi Hubei, Tiongkok, pada Desember 2019. Dalam beberapa bulan saja, penyebaran penyakit ini telah menyebar ke berbagai negara, baik di Asia, Amerika, Eropa, dan Timur Tengah serta Afrika. Pada tanggal 11 Maret 2020, Organisasi Kesehatan Dunia atau World Health Organization (WHO) mendeklarasikan penyebaran COVID-19 dikategorikan sebagai pandemi. Pandemi virus corona (covid-19) yang sudah melanda berbagai negara termasuk Indonesia, Pertama kali mengkonfirmasi kasus covid-19 di Indonesia pada Senin 2 Maret 2020. Penyebaran virus yang sangat cepat pada bulan-bulan setelahnya, banyak yang menghubungkan penyebab lonjakan jumlah positif covid-19 tersebut dengan kondisi yang berbeda di setiap kabupaten/kota, baik dari sisi jumlah penduduk, ketaatan masyarakat, kebijakan pemerintahan, penanganan dan jumlah pasar serta masih ada beberapa hal lain yang menjadi pertimbangan penilaian terhadap terus bertambahnya wabah virus tersebut pada setiap bulannya. Dari analisa yang dilakukan telah didapatkan 3 pengelompokan data (cluster), dimana data ini diolah menggunakan pemrograman python.

Kata kunci : *Data Mining, Covid-19, K-Means, Python,*

tenaga medis terhadap penyebaran kasus covid-19.

1. PENDAHULUAN

Salah satu yang terkena dampak virus corona adalah kabupaten Malang, dimana kabupaten Malang yang berhubungan langsung dengan kabupaten Pasuruan dan jalur menuju kota Surabaya, dimana batas antar kabupaten dan jalur utama menuju kota Surabaya yang menimbulkan dampak yang sangat signifikan terhadap penyebaran positif kasus virus corona, terutama di kecamatan Singosari.

Penelitian ini bertujuan untuk menganalisa faktor-faktor yang menjadi beberapa penyebab lonjakan kasus positif virus corona yang begitu cepat dengan metode Clustering K-Means dan Regresi Linear yang diimplementasikan pada program R dan QGIS. Penelitian ini menghasilkan informasi mengenai pengaruh jumlah penduduk dan jumlah tenaga medis terhadap perkembangan virus covid-19. Dengan demikian diharapkan bahwa informasi tersebut dapat dijadikan salah satu dasar pengambilan kebijakan terhadap langkah-langkah penanganan terhadap kasus covid-19.

Pada penelitian ini, menggunakan Algoritma K-Means Clustering yang digunakan untuk mendapatkan hasil analisa pengaruh jumlah

2. KAJIAN PUSTAKA

Dalam bagian ini akan dijelaskan mengenai beberapa definisi dan kajian teori yang terkait dengan tujuan penelitian

2.1 Data Mining

Data mining mengolah data mentah yang tersimpan di dalam basis data sehingga menghasilkan informasi yang berguna. Data mentah merupakan data yang disimpan hanya sebagai dokumentasi. Pengolahan data mentah dapat dimanfaatkan untuk pembuatan laporan, mengontrol kebutuhan, menemukan informasi dan pengetahuan baru. Data mining menelusuri data pada database untuk membangun model dan menggunakannya untuk mengenali pola data lain yang tidak tersimpan dalam basis data (Kasih, 2019).

2.2 Python

Python merupakan Bahasa pemrograman berarab-tinggi yang diciptakan oleh Guido Van Rossum pada tahun 1989 di Amsterdam, Belanda. Sebagai Bahasa berarab-tinggi, python menawarkan berbagai kemudahan menulis

program.[16]Seiring dengan kecenderungan pembeli pemrograman berorientasi objek dewasaini, python juga sangat tepat digunakan, mengingat python memang Bahasa pemrograman yang berorientasi objek. Oleh karena itu, keistimewaan tentang 26 pewarisan dan instansi yang ditawarkan pada Bahasa yang berorientasi pada objek juga dapat diwujudkan pada python, dengan kata lain, python mendukung konsep 29 reusability(suatu kemudahan untuk mengembangkan kode terhadap kode yang sudah tersedia) (Kadir, 2005)

2.3 COVID-19

COVID-19 adalah penyakit yang menular,dan ditandai oleh gejala pada bagian pernapasan akut (SARS-CoV-2). Virus ini merupakan keluarga besar Coronavirus yang dapat menyerang hewan. Penularan virus ini terjadi jika adanya kontak antar sesama manusia. Ketika menyerang manusia, Virus ini dapat menyebabkan penyakit infeksi saluran pernafasan, seperti MERS (Middle East Respiratory Syndrome), dan SARS (Severe Acute Respiratory Syndrome).[1] Virus ini telah tersebar lebih dari 190 negara lainnya.Oleh karena itu COVID-19 ditetapkan sebagai pandemi oleh WHO. Sampai tanggal 8 Mei 2020, telah terjadi 3.672.238 kasus dan 245.045 kasus kematian di dunia.Sementara di Indonesia terdapat 12.776 kasus positif COVID-19 dan 930 kasus kematian. Dikarenakan pandemi global yang terjadi banyak pihak berupaya ikut berperan serta dalam mengatasi. Para dokter umum dan spesialis angkat bicara Bersama guna memberi penjelasan singkat kepada masyarakat maupun imbauan agar menjaga kebersihan diri dan lingkungan sekaligus tak banyak keluar rumah.

2.4 K-Means

K-means clustering merupakan salah satu metode klaster analisis non hirarki yang berusaha untuk mempartisi objek yang ada kedalam satu atau lebih klaster atau kelompok objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu klaster yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan kedalam klaster yang lain.

Langkah-langkah algoritma K-Means adalah sebagai berikut:

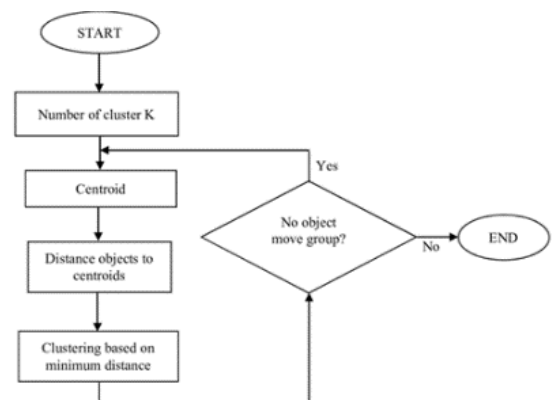
- Tentukan jumlah cluster
- Alokasikan data kedalam cluster secara random
- Hitung centroid/rata-rata dari data yang ada di masing-masing cluster dan Alokasikan masing-masing data ke centroid/rata-rata terdekat berikut:

$$D(a, b) = \sqrt{\sum_{i=1}^n (b_i - a_i)^2}$$

- Kembali ke langkah (c), apabila masih ada data yang berpindah cluster atau apabila ada perubahan nilai centroid.

3. METODOLOGI

Pada bagian ini akan dibahas metode serta langkah- langkah implementasi dalam Analisa pengaruh jumlah tenaga medis. Tahapan penelitian yang dilakukan mulai dari pengumpulan data, studi literatur, implementasi, pengujian dan penarikan kesimpulan.



Gambar 3.1 Flowchart Clustering K-Means

Sebagaimana yang telah dijelaskan pada gambar 3.1 untuk menentukan jumlah cluster dan titik pusat metode clustering k-means membutuhkan ketepatan data yang akan digunakan, setelah itu setiap titik pusat pada masing-masing cluster akan membuat kelompok-kelompok sesuai dengan jarak kedekatan tiap data dengan pusat cluster. Setelah melakukan pengelompokan maka akan dihasilkan data yang sudah diolah dari clustering tersebut

4. HASIL DAN PEMBAHASAN

Pembahasan dan hasil pengujian yang berhubungan dengan data penelitian yang sudah diolah (Data positif kasus covid-19, data jumlah tenaga medis). untuk melakukan pengujian menggunakan data yang sudah ada dengan menggunakan Bahasa pemrograman Python.

4.1 Pengumpulan Data

Pada penelitian ini data di kumpulkan dari jumlah hasil positif covid-19 yang terdapat pada 33 kecamatan di kabupaten malang dimana data terakhir diambil pada tanggal 14 September 2020, dimana data ini merupakan data penyebaran covid-19 di daerah kabupaten malang

LAPORAN COVID-19 KABUPATEN MALANG													
TANGGAL : 14 September 2020													
N O	KECAMATAN	PROBAB LE	SUSPEK			JUMLAH KASUS				KONTAK TERSEKUTU GEDUNG OBSERV	SEMBUR	MENINGG AL	JUM LAH
			HAWAT RS	GEDUNG OBSERV	ISOLASI MANDIR	DISCAR DED	JUMLAH	DIRAWA T RS	ISOLA SI				
1	AMPELGADING		2	2	2	9	13					4	3
2	BANTUR		6	2	10	18					1	1	2
3	BULULAWANG		7	7	13	28	1	2				29	31
4	DAMPIT		4	2	14	20	1					14	15
5	DAU		9	6	47	62		4				31	36
6	DONOMULYO		5		2	7						10	10
7	GEDANGAN			2	4	6							4
8	GONDANGLEGI		15		3	9	21		1			7	2
9	JABUNG		7	11	17	25		1	1			14	16
10	KALIPARE		1	1	7	9			1			7	8
11	KARANGPLOSO		28	2	44	74	1					65	4
12	KASEMBON		2		1	3							3
13	KEPANJEN		26	1	14	56	37	5				41	2
14	KROMENGAN		3	2	9	14			1			8	1
15	LAWANG		27	12	36	135	2	4				30	10
16	NGAJUM		12	2	9	23						9	13
17	NGANTANG		2	1	1	10	21	1	1			3	2
18	PAGAK			6	3	5	8					4	4
19	PAGELARAN		6	1	2	13	22			1	1	6	1
20	PAKIS		1	22	10	62	34	5	6			63	3
21	PAKISAJI		1	9	6	23	38	1	5			39	3
22	PONCOKUSUMO		1	3	6	12	27	1				6	7
23	PULOH		5	1	35	42						11	2
24	SBR MJC VETAM		4		3	6	13					8	8
25	SINGOSARI		1	51	10	226	207	1	5			182	5
26	SUMBERPOCUNG		1		3	9	18	2				5	7
27	TAJINAN		1	6	3	21	30		2	1		2	1
28	TIRTOTUDO				1	5	6					1	1

Gambar 4.1 Data positif covid-19 kabupaten Malang

Data dibawah ini merupakan data yang sudah di filter berdasarkan keterkaitan jumlah penduduk, tenaga medis dan positif covid berdasarkan setiap kecamatan yang berada pada kabupaten malang. Dimana data ini akan digunakan untuk pengelompokan data menggunakan metode clustering k-means

NO	KECAMATAN	POSITIF	TENAGA_MEDIS	JUMLAH_PENDUDUK
1	AMPELGADING	4	44	54456
2	BANTUR	2	28	33549
3	BULULAWANG	31	42	70675
4	DAMPIT	15	52	66084
5	DAU	36	44	69912
6	DONOMULYO	10	56	63777
7	GEDANGAN	4	39	53652
8	GONDANGLEGI	10	34	48183
9	JABUNG	16	50	73896
10	KALIPARE	8	37	62581
11	KARANGPLOSO	50	43	77546
12	KASEMBON	0	32	31133
13	KEPANJEN	48	55	105960
14	KROMENGAN	10	26	39132
15	LAWANG	106	83	106631
16	NGAJUM	8	39	49657
17	NGANTANG	12	34	57769
18	PAGAK	4	28	29129
19	PAGELARAN	9	32	67659
20	PAKIS	83	51	138849
21	PAKISAJI	48	50	85049
22	PONCOKUSUMO	7	50	94792

Gambar 4.2 Data positif covid-19, Tenaga Medis dan Jumlah Penduduk

4.2 Penelitian

Dari data yang sudah di kumpulkan maka adan dilakukan langkah awal dalam melakukan penelitian dengan menggunakan klastering atau pengelompokan data berdasarkan klaster yang sudah ditentukan menggunakan K-Means.

4.3 Pengelompokan Data (Clustering)

Pada langkah ini dilakukan perhitungan K-Means menggunakan bahasa pemrogram Python dengan data jumlah positif covid-19 kabupaten malang dengan 33 kecamatan yang sudah dikumpulkan.

Berikut Syntax clustering K-Means pada Bahasa pemrograman Python;

Proses upload dataset kedalam program python

```
import seaborn as sns
import pandas as pd
from google.colab import files
clusteringbaru=files.upload()

Choose Files clusteringbaru.csv
• clusteringbaru.csv(application/vnd.ms-excel) - 853 bytes, last modified: 12/20/2020 - 100% done
Saving clusteringbaru.csv to clusteringbaru.csv

import matplotlib.pyplot as plt
import numpy as np
from sklearn.cluster import KMeans
from sklearn.preprocessing import MinMaxScaler

KabMalang = pd.read_csv('clusteringbaru.csv')
KabMalang.head(20)
#KabMalang.info()
```

Dataset yang telah di tampilkan

	NO	KECAMATAN	POSITIF	TENAGA_MEDIS	JUMLAH_PENDUDUK
0	1	AMPELGADING	4	44	54456
1	2	BANTUR	2	28	33549
2	3	BULULAWANG	31	42	70675
3	4	DAMPIT	15	52	66084
4	5	DAU	36	44	69912
5	6	DONOMULYO	10	56	63777
6	7	GEDANGAN	4	39	53652
7	8	GONDANGLEGI	10	34	48183
8	9	JABUNG	16	50	73896
9	10	KALIPARE	8	37	62581
10	11	KARANGPLOSO	50	43	77546
11	12	KASEMBON	0	32	31133
12	13	KEPANJEN	48	55	105960
13	14	KROMENGAN	10	26	39132
14	15	LAWANG	106	83	106631
15	16	NGAJUM	8	39	49657
16	17	NGANTANG	12	34	57769
17	18	PAGAK	4	28	29129
18	19	PAGELARAN	9	32	67659
19	20	PAKIS	83	51	138849

Menghapus dua kolom dan menyimpan data ke dataframe baru

```
#Menghapus dua kolom dan menyimpan data ke DataFrame baru
KabMalang = KabMalang.drop(['NO', 'KECAMATAN', 'JUMLAH_PENDUDUK'], axis=1)
KabMalang.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 33 entries, 0 to 32
Data columns (total 2 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   POSITIF      33 non-null     int64
1   TENAGA_MEDIS 33 non-null     int64
dtypes: int64(2)
memory usage: 656.0 bytes
```

Melakukan Plotting Data Ke Diagram Scatter

```
#Plotting Data Ke Diagram Scatter
plt.scatter(KabMalang.POSITIF, KabMalang.TENAGA_MEDIS, s=10, c='c', marker='o', alpha=1)
plt.show()
```

Menentukan variable yang akan dikelompokkan

```
#Menentukan variable yang akan dikelompokkan
KabMalang_knn = KabMalang.iloc[:, 0:2]
KabMalang_knn.head(34)
```

Mengubah variable data frame menjadi array sebelum menentukan nilai K.

```
#Mengubah variable data frame menjadi array sebelum menentukan nilai K.
x_array = np.array(KabMalang_knn)
#print(x_array)
```

Menstandarkan kembali ukuran variable array.

```
#Menstandarkan kembali ukuran variable array.
scaler = MinMaxScaler()
x_scaled = scaler.fit_transform(x_array)
```

Menentukan dan konfigurasi fungsi k-means

```
#Menentukan dan konfigurasi fungsi k-means
kmeans = KMeans(n_clusters=3, random_state=123)
```

Menentukan nilai K (cluster) dari data

```
#Menentukan nilai K (cluster) dari data
kmeans.fit(x_scaled)
```

Menambahkan kolom cluster dalam data frame

```
#Menambahkan kolom cluster dalam data frame
KabMalang_knn['kluster'] = kmeans.labels_
print(KabMalang_knn)
```

Memvisualisasikan hasil cluster

```
#Memvisualisasikan hasil cluster
output = plt.scatter(x_scaled[:,0], x_scaled[:,1], s=100, c=KabMalang_knn.kluster, marker='o', alpha=1)
centers = kmeans.cluster_centers_
plt.scatter(centers[:,0], centers[:,1], c='red', s=200, alpha=1, marker='o')
plt.title('Hasil Klustering K-Means')
plt.colorbar(output)
plt.show()
```

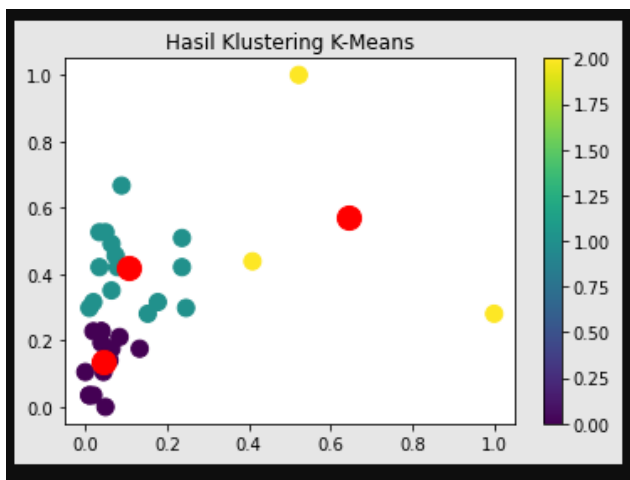
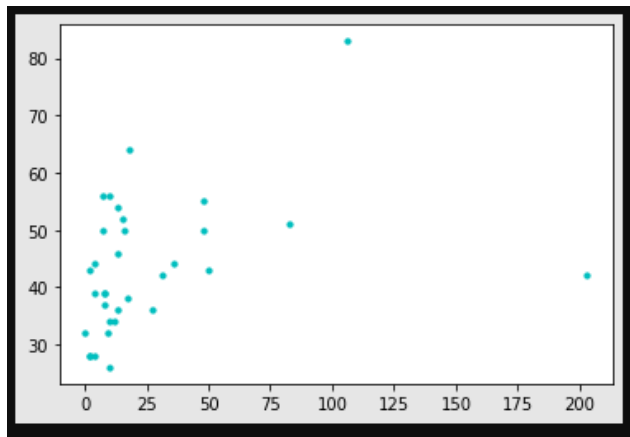
setelah melakukan klastering data maka didapatkan 3 kluster data dari 33 Kecamatan di Kabupaten Malang dimana hasilnya ada pada gambar 4.3 dimana kluter 1 berjumlah 15 daerah, kemudian kluter 2 berjumlah 3 kecamatan dan kluster 3 disini digambarkan denan 0 dimana data dari kluster tersebut didaptnkan 15 daerah, dan hasil sebaran positif covid – 19 pada kabupaten malang didapatkan pada gambar 4.4.

	POSITIF	TENAGA_MEDIS	kluster
0	4	44	1
1	2	28	0
2	31	42	1
3	15	52	1
4	36	44	1
5	10	56	1
6	4	39	0
7	10	34	0
8	16	50	1
9	8	37	0
10	50	43	1
11	0	32	0
12	48	55	1
13	10	26	0
14	106	83	2
15	8	39	0
16	12	34	0
17	4	28	0
18	9	32	0
19	83	51	2
20	48	50	1
21	7	50	1
22	13	36	0
23	8	39	0
24	203	42	2
25	7	56	1
26	17	38	0
27	2	43	1
28	18	64	1
29	13	54	1
30	27	36	0
31	13	46	1
32	2	28	0

Gambar 4.3 Hasil klastering pada pemrograman python

4.4 Grafik Clustering

Setelah mendapatkan hasil clustering dari data covid diatas maka data tersebut dijadikan berupa tabel maka pada tahap ini akan ditampilkan grafik Clustering persebaran data hasil clustering tersebut.



Gambar 4.4 Grafik Clustering Positif Covid – 19
Kabupaten Malang

5. KESIMPULAN

Berdasarkan penelitian diatas maka dapat diambil kesimpulan bahwa jumlah positif covid dengan tenaga medis sangat berpengaruh terhadap dampak penyebar virus corona, Setelah dilakukan pengambilan data dan penelitian guna menentukan klaster atau pengelompokan data persebaran kasus positif covid-19 pada kabupaten Malang di dapat kesimpulan sebagai berikut; Dari 32 kecamatan pada kabupaten Malang didapat 15 kecamatan dengan cluster 1, kemudian 15 kecamatan terdapat pada cluster 3 atau 0 dan 3 kecamatan lainnya pada cluster 2. Dengan adanya hasil clustering tersebut pemerintah daerah maupun masyarakat dapat melakukan kebijakan maupun pengurangan terhadap faktor-faktor yang mempengaruhi penyebaran kasus positif covid-19 di kabupaten Malang.

6. DAFTAR PUSTAKA

- [1] Nango, Dwi Novianti 2012. Penerapan Algoritma K-means untuk Clustering , <http://sro.web.id>.
 - [2] Ken 2009. Clustering Analysis, Part II: K-Means Clustering. <http://www.centerspace.net>.
 - [3] pythonindo.com, syntax dasar python, <https://www.pythonindo.com/sintaks-dasar-python/>
 - [4] Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. J Autoimmun. 2020; published online March 3. DOI: 10.1016/j.jaut.2020.102433
 - [5] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patents infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395(10223):497-506
- A.R. Setiawan, Lembar Kegiatan Literasi Saintifik untuk Pembelajaran Jarak Jauh Topik Penyakit Coronavirus 2019 (COVID-19), Edukatif: Jurnal Ilmu Pendidikan Vol 2, No 1, 2020.

