

Penerapan Metode K-Means pada Data Penduduk Miskin Per Kecamatan Kabupaten Blitar

Dede Nor Alfiansyah^{*1}, Vinna Rahmayanti Setyaning Nastiti², Nur Hayatin³

^{1,2,3}Universitas Muhammadiyah Malang

dedenoralfiansyah@gmail.com^{*1}, vinna.nastiti@gmail.com², noorhayatin@gmail.com³

Abstrak

Kemiskinan adalah kondisi ketidakmampuan pendapatan seseorang dalam pemenuhan kebutuhan pokok hidup sehingga tidak mampu menjamin kelangsungan hidupnya. Kebutuhan hidup meliputi kebutuhan dasar seperti sandang, pangan, tempat tinggal, dan pendidikan. Menurut BPS, yang dikatakan penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran perkapita perbulan dibawah garis kemiskinan. Kabupaten Blitar adalah salah satu kabupaten di provinsi Jawa Timur. Angka kemiskinan di kabupaten Blitar mengalami penurunan sejak tahun 2016 hingga tahun 2019 dan mengalami kenaikan di tahun 2020. Langkah untuk mengetahui angka kemiskinan dihitung berdasarkan sampel survei analisis tingkat kemiskinan dalam survey Sosial – Ekonomi Nasional (Susenas) setiap 2 tahun sekali. Pengelompokan kemiskinan perlu dilakukan setiap tahunnya agar pemerintah dapat mengetahui langkah pencegahan agar kemiskinan tidak meningkat ataupun pemecahan masalah kemiskinan. Pengelompokan kemiskinan dapat dilakukan dengan menggunakan berbagai metode salah satunya adanya dengan menggunakan metode K-Means. Penggunaan metode K-Means ini bertujuan mengelompokkan data per kecamatan kabupaten Blitar yang tergolong penduduk miskin.

Kata Kunci: Kemiskinan, Kabupaten Blitar, K-Means

Abstract

Poverty is the inability of a person's income to fulfill the basic needs of life so that he is unable to guarantee his survival. The necessities of life include basic needs such as clothing, food, shelter, and education. According to BPS, what is said to be the poor are people who have an average monthly per capita expenditure below the poverty line. Blitar Regency is one of the regencies in East Java province. The poverty rate in Blitar district has decreased from 2016 to 2019 and has increased in 2020. The step to find out the poverty rate is calculated based on a sample of the poverty level analysis survey in the National Social-Economic Survey (Susenas) every 2 years. Poverty grouping needs to be done every year so that the government can find out preventive measures so that poverty does not increase or solve the problem of poverty. Poverty grouping can be done using various methods, one of which is the K-Means method. The use of the K-Means method aims to classify data per sub-district of Blitar Regency which is classified as poor.

Keywords: Poverty, Blitar Regency, K-Means

1. Pendahuluan

Kemiskinan adalah karakteristik variasi daerah. Kemiskinan dapat dilihat berdasarkan ketidakmampuan seseorang dalam segi perekonomian untuk memenuhi kebutuhan dasar seperti makanan maupun bukan makanan yang diukur berdasarkan pengeluaran [13]. Faktor seperti sering terjadinya bencana alam, letak geografis, akses pendidikan dan kesehatan, pembangunan infrastruktur, peluang ketenagakerjaan adalah sebagian penyebab kemiskinan [1]. Populasi penduduk di Kabupaten Blitar sejak tahun 2018 hingga 2020 terus mengalami kenaikan. Tahun 2019 jumlah penduduk laki-laki di kabupaten Blitar sebanyak 582.481 orang dan penduduk perempuan sebanyak 350.287 orang. Kepadatan penduduk ini merupakan suatu masalah yang banyak menimbulkan dampak bagi kehidupan sosial ekonomi masyarakat. Dampak yang ditimbulkan seperti sulitnya mencari peluang kerja, pengangguran meningkat, dan segala jenis kebutuhan fasilitas sosial juga meningkat.

Kabupaten Blitar merupakan salah satu kabupaten yang berada di provinsi Jawa Timur. Pusat pemerintahannya terletak di kecamatan Kanigoro. Kabupaten Blitar memiliki luas wilayah sebesar 1.588,79 km². Secara administratif, kabupaten Blitar memiliki 22 kecamatan dengan 248 desa/ kelurahan. Adapun 22 kecamatan yang ada di kabupaten Blitar adalah Panggunrejo, Bakung, Sutojayan, Kademangan, Kanigoro, Garum, Talun, Selopuro, Wlingi, Kesamben, Selorejo, Doko, Nglegok, Sanankulon, Ponggok, Srengat, Wonodadi, Wonotirto, Binangun, Gandusari, Wates, dan Udanawu. Tingkat kemiskinan bersandarkan survei badan pusat statistik (BPS) tahun 2019 di Kabupaten Blitar mencapai 103,75 ribu jiwa (8.94%) [3]. Menurut BPS yang dimaksud penduduk miskin adalah orang yang kemampuan ekonominya berada diambang garis kemiskinan. Jenis garis kemiskinan ada dua yaitu garis kemiskinan makanan (G K M) dan garis kemiskinan non makan (G K N M) yang menjadi indikator pengukuran tingkat kemiskinan. GKM merupakan nilai terkecil untuk kebutuhan bahan makanan. Sedangkan GKNNM merupakan kebutuhan seperti hunian, pakaian, makanan, pendidikan, dan kesehatan. Kategori penduduk miskin dapat diketahui berdasarkan variable kemiskinan yang meliputi kepemilikan kekayaan, hewan ternak, status perkawinan, jenjang pendidikan kepala keluarga dan pasangan, kebutuhan pangan dan kesehatan [2].

Badan pusat statiska (BPS) menganalisa tingkat kemiskinan yang bersandarkan hasil survey Sosial – Ekonomi (Susenas) yang dilakukan tiap dua tahun sekali. Hasil pengukuran kemiskinan diambil berdasarkan sampel survey kurang merata karena tidak menghasilkan suatu ukuran pada tingkat terendah seperti kelurahan. Pengumpulan data diambil sesuai blok sensus. Blok sensus adalah daerah kerja bagi petugas pecacah survey BPS.

Salah satu permasalahan kemiskinan yang belum terselesaikan adalah hasil survey yang diberikan pemerintah tidak sesuai dengan kebutuhan penduduk. Kesulitan dalam proses penanganan kemiskinan oleh pemerintah adalah program bantuan sosial yang ditargetkan pemerintah kurang tepat sasaran. Hal ini disebabkan karena validasi data yang sering diabaikan sehingga sering menimbulkan data yang kurang akurat. Pengumpulan dan pengukuran data ini mengundang peneliti untuk mengkaji data secara lebih akurat agar pelaksanaan program bantuan sosial yang digunakan untuk mengatasi kemiskinan merata, berguna, dan tepat sasaran. Penulis menerapkan metode Algoritma *K-Means Clustering*. Penggunaan metode *K-Means* ini bertujuan mengelompokkan data per kecamatan kabupaten Blitar yang tergolong penduduk miskin.

Prosedur *K-Means* dilakukan dengan cara: (1) penetapan kelompok. (2) mengalokasikan data pada kelompok. (3) menghitung *centroid* disetiap kelompok yang sudah ditentukan. Letak *centroid* diperoleh dari nilai tengah (mean) keseluruhan data perkelompok. (4) peruntukkan tiap - tiap data ke *centroid* terdekat. Menghitung jarak data ke tiap titik pusat *cluster* menggunakan teori jarak *Euclidean* dengan Persamaan 1.

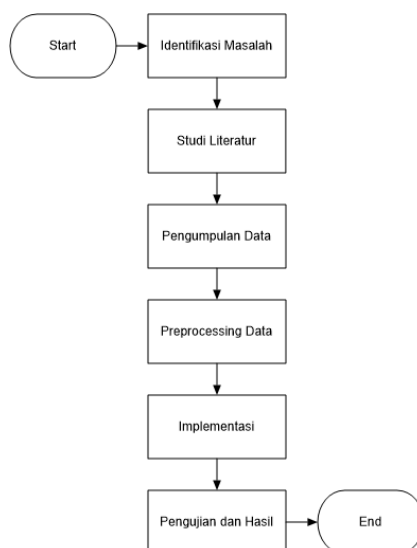
$$De = \sqrt{(xi - si)^2 + (yi - ti)^2} \quad (1)$$

Pengalokasikan ulang data pada tiap – tiap kelompok dalam *K-Means* yang bersandarkan perbandingan jarak antara data ke *centroid* pada kelompok [11].

Penelitian terdahulu yang relevan adalah penelitian yang dilakukan oleh Nanda dan Faisal (2018) dengan judul Penerapan Metode *K-Means* Pemetaan Calon Penerima JAMKESDA. dari penelitian tersebut dihasilkan bahwa metode *K-Means* dapat diterapkan untuk pemetaan masyarakat berdasarkan tingkat kemiskinan di kelurahan Kemuning, pada proses pengelompokkan sampel data yang digunakan sebanyak 440 dan didapatkan 2 *cluster* yang divalidasi menggunakan *Davies Bouldin Index* (DBI). berdasarkan hasil pengelompokkan tersebut sebanyak 334 anggota termasuk kedalam kelompok idak mampu dan 106 anggota kelompok mampu. Penelitian terdahulu yang relevan berikutnya adalah penelitian yang dilakukan oleh Nayuni, dkk (2020). Penelitian tersebut berjudul “Penerapan Algoritma *K-Means* dalam Menentukan Tingkat Penyebaran Pandemi *Covid-19* di Indonesia”. Berdasarkan penelitian tersebut didapatkan hasil bahwa algoritma *K-means* dapat menentukan pusat *cluster* yaitu *cluster* pertama berada pada provinsi DKI Jakarta, *cluster* kedua provinsi Jawa Barat, dan *cluster* ketiga provinsi Banten. Pusat *cluster* pertama adalah daerah penyebaran *Covid 19* terbesar di Indonesia dengan kasus positif dan meninggal terbanyak.

2. Metode Penelitian

Langkah-langkah pada penelitian ini tersusun sesuai Gambar 1 berikut:



Gambar 1. Rancangan Penelitian

- Identifikasi Masalah

Langkah awal yang dilakukan dalam penelitian ini adalah mengidentifikasi masalah sebagai landasan awal untuk memahami dan menentukan solusi penyelesaian serta metode yang akan digunakan. Berdasarkan uraian yang sudah dijelaskan pada latar belakang, maka penulis memutuskan untuk mengangkat permasalahan kemiskinan di Kabupaten Blitar.

Salah satu permasalahan kemiskinan yang belum terselesaikan adalah hasil survey yang diberikan pemerintah tidak sesuai dengan kebutuhan penduduk. Kesulitan dalam proses penanganan kemiskinan oleh pemerintah adalah pembagian bantuan sosial yang tidak merata dan tidak tepat sasaran. Hal ini disebabkan karena validasi data yang sering diabaikan sehingga sering menimbulkan data yang kurang akurat. Penulis akan melakukan pendugaan/prediksi penduduk yang tergolong miskin dengan menghitung data yang diperoleh dari BPS dan PLN menggunakan metode *K-Means*. Penelitian ini diharapkan dapat menghasilkan klasterisasi penduduk yang tergolong miskin, sehingga pemerintah dapat mengetahui jumlah penduduk miskin per kecamatan di kabupaten Blitar. Hasil pengelompokan ini diharapkan menjadi bahan rujukan sebagai fokus untuk mengurangi permasalahan kemiskinan dan pemecahan masalah kemiskinan melalui program pemerintah seperti bantuan sosial dapat di distribusikan kepada penduduk miskin dengan baik dan tepat sasaran.

- Studi Literatur

Pada tahap ini akan dilakukan pemahaman tentang konsep dasar algoritma – algoritma yang digunakan dalam penelitian yang akan dilakukan. Adapun literatur yang digunakan adalah jurnal dan buku yang memiliki tema atau bahasan yang sesuai dengan penelitian yang akan dilakukan, yaitu tentang kemiskinan dan metode algoritma *K-Means*.

- Pengumpulan Data

Pada penelitian ini, data yang diperlukan berupa data yang diambil dari BPS dan PLN Kabupaten Blitar. Data diperoleh dari *website* resmi BPS Kabupaten Blitar. Data dari PLN diperoleh secara langsung dari kantor PLN Kabupaten Blitar.

- *Preprocessing Data*

Data yang telah didapatkan kemudian diolah sebelum memasuki proses klasifikasi. Pada tahap *preprocessing* ini, data mentah yang telah berhasil diambil pada tahapan sebelumnya akan diolah terlebih dahulu dikarenakan data yang diterima masih acak dan perlu di kelompokkan berdasarkan kebutuhan data yang akan diolah. Pada tahap ini dilakukan penentuan kolom yang akan dijadikan sebagai acuan dalam klasterisasi, pada penelitian ini kolom yang digunakan

sebagai acuan untuk menentukan kluster tingkat kemiskinan kecamatan menggunakan kolom, "JKM", "PSS", "PLN", "KP", "SP".

Preprocessing berfungsi untuk mengeliminasi atau menghilangkan data yang tidak sesuai atau secara garis besar mengubah data tersebut agar lebih mudah diproses oleh system dengan kata lain tahapan ini adalah tahapan awal untuk menyetarakan data yang sudah berhasil dikumpulkan.

- Implementasi

Dari tahapan awal mengidentifikasi masalah, studi literatur, pengumpulan data, hingga *preprocessing*, kemudian data yang sudah melewati tahapan – tahapan sebelumnya akan diimplementasikan dengan metode *K-Means*.

- Pengujian dan Hasil

Proses pengujian pada penelitian ini akan dilakukan klasterisasi data dengan menggunakan tiga kluster yang dilabeli menggunakan "0", "1". Label tersebut akan menunjukkan nilai terendah sampai nilai tertinggi.

Hasil dari penelitian ini nantinya berupa hasil klasterisasi data akan digabungkan dengan variable-variabel kemiskinan dan akan diketahui hasil klasifikasi kluster berdasarkan jenis label yang diperoleh. Hasil klasterisasi kemudian divalidiasi menggunakan *silhouette coefficient* untuk mengukur kualitas kluster.

1. Analisa Kebutuhan Sistem

Analisis kebutuhan sistem berisi tentang analisa kebutuhan perangkat keras dan perangkat lunak yang digunakan dalam penelitian ini. Adapun dalam penelitian ini menggunakan kebutuhan seperti pada Tabel 1 dan Tabel 2.

- Kebutuhan perangkat keras (*hardware*)

Tabel 1. Tabel Kebutuhan Perangkat Keras

No	Perangkat	Jumlah	Spesifikasi
1	Laptop	1	Processor : Intel Pentium inside RAM : 2 GB HDD : 320 GB

- Kebutuhan perangkat lunak (*software*)

Tabel 2. Tabel Kebutuhan Perangkat Lunak

No	Software
1	Sistem Operasi Windows 10
2	Python
4	Micorosoft Word
5	Microsoft Excel

3. Hasil Penelitian dan Pembahasan

3.1 Data

Data pada penelitian ini menggunakan data penduduk miskin tiap Kecamatan yang terdiri dari 22 Kecamatan di Kabupaten Blitar. Data tersebut mencakup:

- Kec = Kecamatan
- JKM = Jumlah Keluarga Miskin
- PSS = Penduduk Sedang Sekolah
- PLN = Pelanggan PLN
- KP = Kepadatan Penduduk
- SP = Sektor Pertanian

Data yang diperoleh telah berformat csv, data tersebut dibaca dengan menggunakan library pandas dan disimpan pada variable "data" seperti pada Gambar 2.

```
#RAK = rata-rata anggota keluarga
#JKM = jumlah keluarga miskin
#PSS = penduduk sedang sekolah
#PLN = pelanggan pln
#KP = kepadatan penduduk
#SP = sektor pertanian
data = pd.read_csv("data.csv")
data.head(10)
```

Gambar 2. Source Code Import Data

Pada variable “data” terdapat data keseluruhan dan dapat ditampilkan 10 data teratas menggunakan fungsi “head()” seperti pada Gambar 3 berikut.

Kec	RAK	JKM	PSS	PLN	KP	SP
1	2	416	750	7544	230	251
2	2	419	914	10616	218	288
3	2	495	1034	12192	347	666
4	2	445	615	8324	410	863
5	2	591	844	12843	566	122
6	2	438	1292	14120	1081	1450
7	2	541	1906	19332	622	833
8	2	878	1934	23106	1408	1721
9	2	653	1391	17959	1221	2350
10	2	591	1199	11841	1020	1810

Gambar 3. Sampel Data

3.2 Preprocessing Data

Selanjutnya adalah menentukan kolom yang akan dijadikan sebagai acuan dalam klusterisasi, pada penelitian ini kolom yang digunakan sebagai acuan untuk menentukan kluster tingkat kemiskinan kecamatan menggunakan kolom, “JKM”, “PSS”, “PLN”, “KP”, “SP”. Data tersebut akan dipisah dan dibentuk menjadi array dan disimpan dalam variable “clusterdata_array” seperti pada Gambar 4 berikut.

```
: clusterdata = data.loc[:,["JKM","PSS","PLN","KP","SP"]]
: clusterdata_array = np.array(clusterdata)
: print(clusterdata_array)

[[ 416  750  7544  230  251]
 [ 419  914 10616  218  288]
 [ 495 1034 12192  347  666]
 [ 445  615  8324  410  863]
 [ 591  844 12843  566  122]
 [ 438 1292 14120 1081 1450]
 [ 541 1906 19332  622  833]
 [ 878 1934 23106 1408 1721]
 [ 653 1391 17959 1221 2350]
 [ 591 1199 11841 1020 1810]
 [ 583 1233 14580  867 1758]
 [ 437  583 10458  678  979]
 [ 393  749 11319  540 1815]
 [ 582 1401 15075  769 1525]
 [ 602 1827 19992  767 2580]
 [ 759 1869 19414 1205 2165]
 [ 763 2103 20759  759 1510]
 [ 819 1180 16736 1700 1207]
 [ 826 2456 30239  986 2033]
 [ 804 1806 19462 1221 1644]
 [ 765 1460 13875 1164 2209]
 [ 783 1121 12139 1003 2222]]
```

Gambar 4. Pembentukan Array Data

Langkah selanjutnya adalah data tersebut akan di normalisasi sebelum dihitung menggunakan metode klusterisasi. Normalisasi data dilakukan menggunakan “*MinMaxScaler()*” yang didapatkan dari library “*sklearn.preprocessing*” seperti pada Gambar 5 berikut.

```
from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
clusterdata_normalize = scaler.fit_transform(clusterdata_array)
clusterdata_normalize

array([[0.04742268, 0.08916177, 0.          , 0.00809717, 0.05248169],
       [0.05360825, 0.17672184, 0.13536021, 0.          , 0.06753458],
       [0.21030928, 0.24079018, 0.20480282, 0.08704453, 0.22131814],
       [0.10721649, 0.01708489, 0.03436888, 0.12955466, 0.30146461],
       [0.40824742, 0.13934864, 0.23348755, 0.23481781, 0.          ],
       [0.09278351, 0.37853711, 0.28975545, 0.58232119, 0.54027665],
       [0.30515464, 0.70635344, 0.51940956, 0.27260459, 0.28925956],
       [1.          , 0.72130272, 0.68570177, 0.80296896, 0.65052889],
       [0.53608247, 0.43139349, 0.45891165, 0.67678812, 0.90642799],
       [0.40824742, 0.32888414, 0.18933686, 0.54116059, 0.68673718],
       [0.39175258, 0.34703684, 0.31002423, 0.43792173, 0.66558177],
       [0.09072165, 0.          , 0.12839833, 0.31039136, 0.34865745],
       [0.          , 0.08862787, 0.16633622, 0.21727395, 0.68877136],
       [0.38969072, 0.43673251, 0.33183521, 0.37179487, 0.57078926],
       [0.43092784, 0.66417512, 0.54849086, 0.37044534, 1.          ],
       [0.75463918, 0.68659904, 0.52302269, 0.66599199, 0.83116355],
       [0.76288666, 0.81153233, 0.58228685, 0.36504723, 0.56468674],
       [0.87835052, 0.31873999, 0.40502313, 1.          , 0.44141579],
       [0.89278351, 1.          , 1.          , 0.51821862, 0.77746135],
       [0.84742268, 0.65296316, 0.52513777, 0.67678812, 0.61920266],
       [0.76701031, 0.46823278, 0.27896012, 0.63832659, 0.84906428],
       [0.80412371, 0.28723972, 0.20246755, 0.52968961, 0.85435313]])
```

Gambar 5. Normalisasi Data

Berdasarkan gambar diatas menunjukkan *source code* dan hasil normalisasi data yang disimpan dalam variable “*clusteredata_normalize*”.

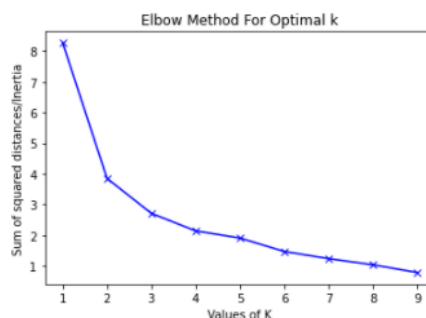
3.3 Klusterisasi Data

Pada klusterisasi data yang pertama dilakukan adalah menguji berapa nilai kluster yang optimal dari sebuah data sebelum dibentuk menjadi sebuah kluster dengan menggunakan metode *elbow*, seperti pada Gambar 6 berikut.

```
Sum_of_squared_distances = []
K = range(1,10)
for num_clusters in K :
    kmeans = KMeans(n_clusters=num_clusters)
    kmeans.fit(clusterdata_normalize)
    Sum_of_squared_distances.append(kmeans.inertia_)
plt.plot(K, Sum_of_squared_distances, 'bx-')
plt.xlabel('Values of K')
plt.ylabel('Sum of squared distances/Inertia')
plt.title('Elbow Method For Optimal k')
plt.show()
```

Gambar 6. Pembentukan Grafik K

Pada gambar diatas percobaan dilakukan dengan menjalankan klusterisasi *K-Means* pada dataset untuk rentan nilai k yang berbeda yaitu 1 sampai 9 untuk setiap k yang akan menghitung jarak rata-rata pada pusat cluster di semua data. Adapun hasil grafik sebagai berikut.



Gambar 7. Grafik K

Pada Gambar 7 diatas dapat dilihat grafik dari masing masing k, berdasarkan gambar tersebut dapat diketahui bahwa nilai k yang paling baik untuk digunakan adalah dalam klasterisasi dengan menggunakan k = 2, karena pada nilai k = 2 terdapat kuadran yang turun secara tiba-tiba hingga membentuk sudut yang lebih dekat dengan siku-siku dibandingkan dengan kuadran yang lain.

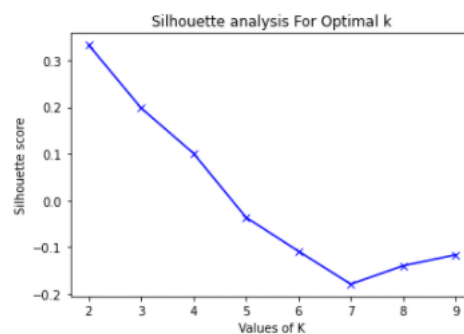
Selanjutnya dilakukan uji coba menggunakan nilai *silhouette coefficient* pada setiap k untuk menentukan nilai k, *silhouette coefficient* yang akan menentukan kemiripan satu titik data dalam cluster dibandingkan dengan cluster lain, Adapun source code perhitungan *silhouette coefficient* seperti pada Gambar 8 berikut.

```
range_n_clusters = [2, 3, 4, 5, 6, 7, 8, 9]
silhouette_avg = []
for num_clusters in range_n_clusters:
    # initialise kmeans
    kmeans = KMeans(n_clusters=num_clusters)
    kmeans.fit(clusterdata_normalize)
    cluster_labels = kmeans.labels_

    # silhouette score
    silhouette_avg.append(silhouette_score(data, cluster_labels))
plt.plot(range_n_clusters, silhouette_avg, 'bx-')
plt.xlabel('Values of K')
plt.ylabel('Silhouette score')
plt.title('Silhouette analysis For Optimal k')
plt.show()
```

Gambar 8. Source Code Silhouette Coefficient

Pada gambar diatas menunjukkan pengujian *silhouette coefficient* pada rentan nilai 2 sampai 9 pada setiap nilai k dan menghasilkan sebuah plot pada grafik seperti pada Gambar 9 berikut.



Gambar 9. Grafik Silhouette Coefficient

Pada gambar diatas menunjukkan hasil grafik dari nilai *silhouette coefficient*, nilai rentan *silhouette coefficient* berada diantara [-0.2, 0.33]. pada pengujian *silhouette coefficient* nilai yang mendekati satu menunjukkan titik terbaik atau nilai k terbaik untuk digunakan dalam klaterisasi. Berdasarkan grafik tersebut nilai k yang baik berada pada k = 2 karena nilai *silhouette coefficient* nya mendekati nilai 1.

Berdasarkan 2 pengujian diatas dapat disimpulkan nilai k yang baik digunakan dalam clusterisasi data adalah dengan menggunakan nilai k = 2, selanjutnya klasterisasi data dilakukan pada data dengan nilai k = 2 seperti pada Gambar 10 berikut.

```
kmeans_sel = KMeans(init='k-means++', n_clusters=2, n_init=100,
                    random_state=100).fit(clusterdata_normalize)
labels = pd.DataFrame(kmeans_sel.labels_)
clustered_data = data.assign(Cluster=labels)
```

Gambar 10. Klasterisasi Data

Berdasarkan gambar diatas dapat diketahui bahwa hasil klasterisasi dibentuk kedalam data *frame* dengan menggunakan *library* *pandas* dan kemudian digabungkan kedalam “data” untuk mengetahui data yang masuk kedalam masing-masing klaster. Adapun hasil klasterisasi data sebagai berikut.

	Kec	RAK	JKM	PSS	PLN	KP	SP	Cluster
0	1	2	416	750	7544	230	251	0
1	2	2	419	914	10616	218	288	0
2	3	2	495	1034	12192	347	666	0
3	4	2	445	615	8324	410	863	0
4	5	2	591	844	12843	566	122	0
5	6	2	438	1292	14120	1081	1450	0
6	7	2	541	1906	19332	622	833	0
7	8	2	878	1934	23106	1408	1721	1
8	9	2	653	1391	17959	1221	2350	1
9	10	2	591	1199	11841	1020	1810	0
10	11	2	583	1233	14580	867	1758	0
11	12	2	437	583	10458	678	979	0
12	13	2	393	749	11319	540	1815	0
13	14	2	582	1401	15075	769	1525	0
14	15	2	602	1827	19992	767	2580	1
15	16	2	759	1869	19414	1205	2165	1
16	17	2	763	2103	20759	759	1510	1
17	18	2	819	1180	16736	1700	1207	1
18	19	2	826	2456	30239	986	2033	1
19	20	2	804	1806	19462	1221	1644	1

Gambar 11. Klasterisasi Data

Pada Gambar 11 diatas menunjukkan sebagian data hasil klasterisasi dengan $k = 2$, kolom “Cluster” menunjukkan data tersebut termasuk kedalam klaster 0 atau 1. Langkah selanjutnya untuk membagi data berdasarkan klaster dapat dilakukan dengan menggunakan *source code* berikut.

```
cluster0 = []
cluster1 = []

for i in range(len(clustered_data)):
    if(clustered_data["Cluster"][i] == 0):
        a = clustered_data.loc[i,["Kec", "RAK", "JKM", "PSS", "PLN", "KP", "SP", "Cluster"]]
        cluster0.append(a)
    elif(clustered_data["Cluster"][i] == 1):
        a = clustered_data.loc[i,["Kec", "RAK", "JKM", "PSS", "PLN", "KP", "SP", "Cluster"]]
        cluster1.append(a)

cluster0 = pd.DataFrame(cluster0)
cluster1 = pd.DataFrame(cluster1)
```

Gambar 12. Pembagian Data berdasarkan Klaster

Pada Gambar 12 diatas jika kolom “Cluster” menunjukkan 0 maka data akan dimasukkan kedalam *variable* “cluster0” yang kemudian data tersebut dirubah menjadi dataframe dan menghasilkan bentuk data seperti berikut.

	Kec	RAK	JKM	PSS	PLN	KP	SP	Cluster
0	1	2	416	750	7544	230	251	0
1	2	2	419	914	10616	218	288	0
2	3	2	495	1034	12192	347	666	0
3	4	2	445	615	8324	410	863	0
4	5	2	591	844	12843	566	122	0
5	6	2	438	1292	14120	1081	1450	0
6	7	2	541	1906	19332	622	833	0
9	10	2	591	1199	11841	1020	1810	0
10	11	2	583	1233	14580	867	1758	0
11	12	2	437	583	10458	678	979	0
12	13	2	393	749	11319	540	1815	0
13	14	2	582	1401	15075	769	1525	0

Gambar 13. Klaster ke "0"

Berdasarkan Gambar 13 diatas menunjukkan semua data yang termasuk dalam klaster "0" yaitu data pada kecamatan 1, 2, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14. Berdasarkan gambar diatas yang termasuk kategori penduduk miskin adalah kecamatan – kecamatan berikut: Bakung, Wonotirto, Panggungrejo, Wates, Binangun, Sutojayan, Kademangan, Selopuro, Kesamben, Selorejo, Doko, dan Wlingi.

Kemudian jika kolom "cluster" menunjukkan 1 maka data akan dimasukkan kedalam variable "cluster1" yang kemudian data tersebut dirubah menjadi dataframe dan menghasilkan bentuk data seperti berikut.

	Kec	RAK	JKM	PSS	PLN	KP	SP	Cluster
7	8	2	878	1934	23106	1408	1721	1
8	9	2	653	1391	17959	1221	2350	1
14	15	2	602	1827	19992	767	2580	1
15	16	2	759	1869	19414	1205	2165	1
16	17	2	763	2103	20759	759	1510	1
17	18	2	819	1180	16736	1700	1207	1
18	19	2	826	2456	30239	986	2033	1
19	20	2	804	1806	19462	1221	1644	1
20	21	2	765	1460	13875	1164	2209	1
21	22	2	783	1121	12139	1003	2222	1

Gambar 14. Klaster ke "1"

Pada Gambar 14 diatas menunjukkan semua data yang termasuk dalam klaster "1" yaitu data pada kecamatan 8, 9, 15, 16, 17, 18, 19, 20, 21, 22.

4. Kesimpulan

Berdasarkan hasil penghitungan menggunakan metode *K-Means*, maka diketahui hasil dan dapat disimpulkan sebagai berikut:

1. Pengelompokkan kecamatan – kecamatan dengan penduduk miskin diketahui berdasarkan variabel jumlah keluarga miskin, penduduk sedang sekolah, pelanggan PLN, kepadatan penduduk, dan sektor pertanian.
2. Jumlah sampel yang digunakan sebanyak 22 kecamatan di Kabupaten Blitar dan menghasilkan 2 klaster yaitu klaster 0 dan klaster 1.

3. Metode K-Means dapat digunakan untuk mengelompokkan penduduk miskin per kecamatan di kabupaten Blitar.
4. Berdasarkan hasil klasterisasi, data klaster 0 adalah kecamatan yang dikelompokkan sebagai kecamatan dengan penduduk miskin di Kabupaten Blitar.

5. Daftar Notasi

Contoh penulisan notasi dapat diuraikan dengan keterangan sebagai berikut:

De = Euclidean Distance

I = jumlah objek

(x,y) = Koordinat Object

(s,t) = Koordinat Centroid

Referensi

- [1] M. Y. Darsyah and R. Wasono, "Pendugaan Tingkat Kemiskinan di Kabupaten Sumenep dengan Pendekatan SAE."
- [2] W. Isdijoso and A. Suryahadi, "Penetapan Kriteria dan Variabel Pendataan Penduduk Miskin yang Komprehensif dalam Rangka Perlindungan Penduduk Miskin di Kabupaten / Kota."
- [3] BPS Kabupaten Blitar. "Jumlah dan Persentase Penduduk Miskin dan Garis Kemiskinan Menurut Kabupaten Kota di Jawa Timur." <https://www.bps.go.id/indicator/23/621/1/persentase-penduduk-miskin-menurut-kabupaten-kota.html> (3 Januari 2020).
- [4] Nasution, Irmanita., Windarto, Agus Perdana., Fauzan, M, "Penerapan Algoritma K-Means Dalam Pengelompokan Data Penduduk Miskin Menurut Provinsi". Volume 2, No,2 Desember 2020.
- [5] Suryawati. 2004. Teori Ekonomi Mikro. UPP. AMP YKPN. Yogyakarta: Jarnasy
- [6] Suhartini., Yuliani, Ria, "Penerapan Data Mining untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma K-Means di Dusun Bagik Endep Sukamulia Timur". Vol.4 No. 1 Januari 2021.
- [7] Nurwati, Nunung, "Kemiskinan: Model Pengukuran, Permasalahan dan Alternatif Kebijakan". Vol. 10, No.1 , Januari 2008.
- [8] Nazara, Suahasil. 2007. "Pengentasan Kemiskinan : Pilihan Kebijakan dan Program yang Realitis" Dalam Wara Demografi tahun ke 37. No 4 tahun 2007. Jakarta. Lembaga Demografi Universitas Indonesia. Pengangguran, Kemiskinan, dan Pertumbuhan Ekonomi Indonesia.
- [9] Kemsos. "UU 13 Tahun 2011 - Pusat Penelitian dan Penanganan Fakir Miskin" <https://puslit.kemsos.go.id/upload/aturan/files/f5c497935cc2794803aa5f91c3f2404e.pdf> (15 Desember 2019)
- [10] Aras, Zainul., Sarjono, "Analisis Data Mining untuk Menentukan Kelompok Prioritas Penerima Bantuan Bedah Rumah Menggunakan Metode Clustering K-Means (Studi Kasus: Kantor Kecamatan Bahar Utara)" Vol 1 No.2, Desember 2016.
- [11] Duhita, Windha Mega, "Clustering Menggunakan Metode K-Means untuk Menentukan Status Gizi Balita" Vol. 15, No.2, Desember 2015.
- [12] Anggara, Mario., Sujiani, Herry., Nasution, Helfi, "Pemilihan *Distance Measure pada K-Means Clustering* untuk Pengelompokan Member di Alvaro Fitness" Vol. 1, No. 1, 2016.
- [13] BPS "Kemiskinan dan Penduduk Miskin" <https://www.bps.go.id/subject/23/kemiskinan-dan-ketimpangan.html> (7 Juli 2021)
- [14] BPS "Kemiskinan dan Ketimpangan" <https://jatim.bps.go.id/indicator/12/375/1/jumlah-penduduk-provinsi-jawa-timur.html> (7 Juli 2021)
- [15] Arif Muttaqien, Paradigma Baru Pemberantasan Kemiskinan, Rekonstruksi Arah Pembangunan Menuju Masyarakat Yang Berkeadilan, Terbebaskan Dan Demokratis, (Jakarta; Khanata Pustaka LP3ES Indonesia, 2006), dikutip oleh Istiana Herawati , "Dampak Program Pengentasan Kemiskinan Dampak Program Pengentasan Kemiskinan Di Kabupaten Jayapura", (Jurnal Penelitian dan Evaluasi Pendidikan BP2P3KS Kementerian Sosial RI), h. 146