

**OPTIMASI *K-MEANS CLUSTERING* MENGGUNAKAN METODE
ELBOW UNTUK PENGELOMPOKAN RUMAH TANGGA MISKIN DI
SURABAYA**

SKRIPSI



**UIN SUNAN AMPEL
S U R A B A Y A**

Disusun Oleh
FITRIANINGSIH
H72216030

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA**

2021

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : FITRIANINGSIH

NIM : H72216030

Program Studi : Matematika

Angkatan : 2016

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul " OPTIMASI *K-MEANS CLUSTERING* MENGGUNAKAN METODE *ELBOW* UNTUK PENGELOMPOKAN RUMAH TANGGA MISKIN DI SURABAYA ". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 14 Juni 2021

Yang menyatakan,



FITRIANINGSIH
NIM. H72216030

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

Nama : FITRIANINGSIH

NIM : H72216030

Judul Skripsi : OPTIMASI *K-MEANS CLUSTERING* MENGGUNAKAN
METODE *ELBOW* UNTUK PENGELOMPOKAN RU-
MAH TANGGA MISKIN DI SURABAYA

telah diperiksa dan disetujui untuk diujikan.

Surabaya, 14 Juni 2021

Pembimbing



ARIS FANANI M. Kom

NIP. 198701272014031002

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi oleh

Nama : FITRIANINGSIH
NIM : H72216030
Judul Skripsi : OPTIMASI *K-MEANS CLUSTERING* MENGGUNAKAN
METODE *ELBOW* UNTUK PENGELOMPOKAN RU-
MAH TANGGA MISKIN DI SURABAYA

Telah dipertahankan di depan Tim Penguji
pada tanggal 1 Juli 2021

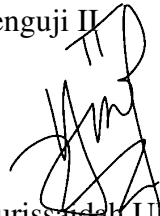
Mengesahkan,
Tim Penguji

Penguji I



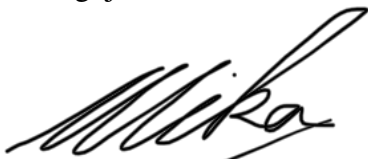
ARIS HANANI M. Kom
NIP. 198701272014031002

Penguji II



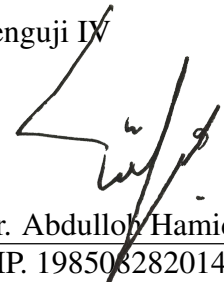
Nurissardah Ulinuha, M. Kom
NIP. 199011022014032004

Penguji III



Wika Dianita Utami, M.Sc
NIP. 199206102018012003

Penguji IV




Dr. Abdulloh Hamid, M.Pd
NIP. 198508282014031003

Mengetahui,

Dekan Fakultas Sains dan Teknologi
UN Sunan Ampel Surabaya




Dr. Hj. Evi Fatmatur Rusyidah, M.Ag
NIP. 197312272005012003



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : Fitrianingsih
NIM : H72216030
Fakultas/Jurusan : Sains dan Teknologi / Matematika
E-mail address : fitrianingsihfitri53@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

Skripsi Tesis Desertasi Lain-lain (.....)
yang berjudul :

OPTIMASI K-MEANS *CLUSTERING* MENGGUNAKAN METODE *ELBOW* UNTUK
.....
PENGELOMPOKAN RUMAH TANGGA MISKIN DI SURABAYA
.....

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara **fulltext** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 10 Agustus 2021

Penulis


(Fitrianingsih)

3.1. Jenis Penelitian	24
3.2. Metode Pengumpulan Data	24
3.3. Variabel Penelitian	25
3.4. Tahap Penelitian	28
IV HASIL DAN PEMBAHASAN	30
4.1. Deskripsi Data	30
4.2. Menentukan Jumlah K <i>Cluster</i> Menggunakan Metode <i>Elbow</i>	30
4.3. Perhitungan Analisis <i>Cluster</i> Menggunakan Metode K-Means	36
4.4. Perhitungan Evaluasi <i>Cluster</i>	42
4.5. Interpretasi <i>Cluster</i>	48
4.5.1. Kategori Kondisi Bangunan	49
4.5.2. Kategori Pangan	51
4.5.3. Kategori Kesehatan	52
4.5.4. Kategori Karakteristik Kepala Rumah Tangga	53
4.5.5. Kategori Kepemilikan Aset	55
4.6. Integrasi Keislaman	57
V PENUTUP	59
5.1. Simpulan	59
5.2. Saran	60
DAFTAR PUSTAKA	61
A SKRIP PROGRAM PYTHON	64

nesia, Tim Nasional Percepatan Penanggulangan Kemiskinan (TNP2K) telah melakukan penetapan Rencana Pembangunan Jangka Menengah Nasional (RPJM) 2020-2024. TNP2K merupakan lembaga yang dibentuk khusus oleh presiden dan diketuai langsung oleh wakil presiden RI untuk meningkatkan penanganan program pengentasan kemiskinan di Indonesia. Adapun target pengurangan kemiskinan yang ditetapkan oleh TNP2K dalam RPJM mencapai 7% target moderat dan 6,5% target optimis pada tahun 2024. Jika penduduk miskin pada 2024 diperkirakan sebanyak 19,75 juta orang, maka dibutuhkan penurunan penduduk miskin sebesar 5,04 juta sampai 6,45 juta orang untuk memenuhi target. Bentuk upaya yang dilakukan oleh TNP2K diantaranya meningkatkan pendapatan dan menurunkan beban pengeluaran masyarakat melalui berbagai perbaikan program bantuan sosial seperti Program Indonesia Pintar Kuliah (PIP-K), Program Keluarga Harapan (PKH), Program Bantuan Pangan (Rastra dan Bantuan Pangan Non-Tunai), Jaminan Kesehatan Nasional (JKN), bantuan subsidi listrik, LPG dan masih banyak bantuan yang diberikan pemerintah (Tim Nasional Percepatan Pengentasan Kemiskinan, 2020).

Berbagai program bantuan yang disediakan pemerintah diberikan kepada rumah tangga yang tergolong kurang mampu, untuk membantu meringankan beban biaya pengeluaran rumah tangga tersebut. Akan tetapi, bantuan yang dibutuhkan setiap rumah tangga yang tergolong tidak mampu memiliki proporsi yang berbeda-beda. Dengan mempertajam ketepatan sasaran dalam menyalurkan berbagai jenis bantuan yang telah direncanakan, dapat membantu pemerintah dalam mencapai target angka pengentasan kemiskinan serta membantu masyarakat kurang mampu mendapatkan hak-hak yang seharusnya mereka peroleh. Allah SWT berfirman dalam Al-Quran Surat An-Nisa ayat 58:

tis, 18% penghasilannya sedikit berkurang dan 14% mengatakan tidak berkurang (Lidwina, 2020) . Tidak hanya mengenai permasalahan tersebut, sebagian warga Surabaya juga masih merasakan adanya bantuan sosial yang tidak tepat sasaran, seperti adanya warga yang masih menerima bantuan padahal orang tersebut sudah meninggal (Andira, 2020) .

Dalam upaya pengentasan kemiskinan sejatinya pemerintah ataupun masyarakat saling memiliki peran penting di dalamnya. Banyak upaya yang dapat dilakukan dalam menghadapi permasalahan kemiskinan, salah satunya yaitu dengan melakukan pengelompokan rumah tangga sesuai dengan kesamaan karakteristik yang dimiliki. Dari pengelompokan rumah tangga tersebut, diharapkan dapat diketahui jenis bantuan apa yang dibutuhkan setiap kelompok. Algoritma K-Means merupakan salah satu metode *clustering* yang dapat digunakan, karena termasuk dalam non-herarki *clustering* sehingga jumlah K *cluster* dapat ditentukan sesuai dengan yang dibutuhkan. Akan tetapi dalam metode K-Means masih terdapat kelemahan yaitu dalam menganalisa jumlah K *cluster* yang baik digunakan. Oleh sebab itu, perlu adanya metode untuk menentukan jumlah K *cluster* yang baik. Salah satu metode yang dapat digunakan yaitu metode *elbow*, kemampuannya dalam menganalisa jumlah K *cluster* dapat digunakan untuk mengoptimalkan kinerja K-Means.

Beberapa penelitian yang pernah dilakukan sebelumnya mengenai klasterisasi rumah tangga miskin dilakukan oleh Dwi Rahayu Utami (2018) dengan judul “Aplikasi Monitoring Keluarga Miskin Menggunakan Metode K-Means *Clustering* Berbasis Mobile GIS (Studi Kasus: PKH Kecamatan Kedungkandang Kota Malang)”, dalam penelitian tersebut menjelaskan mengenai pembuatan aplikasi mobile dengan menggunakan algoritma K-Means *Clustering* untuk mengelompokkan keluarga miskin. Adapun nilai akurasi yang dihasilkan dari pengelompokan sebesar

92,8% atau bisa disebut dengan *Excelent classification* (Utami, 2018).

Penelitian juga dilakukan oleh Nugroho Irawan Febianto dan Nico Dias Palasara (2019) dengan judul penelitian “Analisis *Clustering* K-Means Pada Data Informasi Kemiskinan Di Jawa Barat Tahun 2018”. Pada kesimpulan penelitian tersebut menghasilkan 5 *cluster* yang memetakan wilayah provinsi Jawa Barat berdasar nilai rata-rata tertinggi dan terendah yang diperoleh dari indikator kemiskinan (Febianto dan Nico, 2019).

Carina Kallestal, dkk (2020) melakukan penelitian dengan judul “Assessing the Multiple Dimensions of Poverty. Data Mining Approaches to the 2004-14 Health and Demographic Surveillance System in Cuatro Santos Nicaragua”. Algoritma yang digunakan pada penelitian tersebut yaitu K-Means *cluster*, karena dinilai lebih mudah digunakan dan ditafsirkan. Variabel pada penelitian tersebut dibagi menjadi 5 kelompok, yaitu Kebutuhan dasar yang tidak terpenuhi, aset, kerawanan pangan, intervensi dan variabel individu turunan. Penelitian tersebut menghasilkan 6 *cluster* yang menunjukkan adanya kemiskinan multidimensi. Hasil penelitian tersebut telah dikonfirmasi ketepatannya melalui diskusi yang dilakukan peneliti bersama tokoh masyarakat setempat termasuk tokoh kesehatan, keamanan serta perwakilan dari masyarakat awam (Kallestal et al., 2020).

Penelitian mengenai validitas *cluster* yang dilakukan oleh Khairatia dkk (2019) dengan judul “Kajian Indeks Validitas pada Algoritma K-Means Enhanced dan Kmeans MMCA” yang membandingkan indeks validitas *silhouette index*, *Davies Bouldin Index*, *Dunn* dan *Calinski-Harabasz* menyatakan *silhouette index*, *Davies Bouldin Index* dan *Calinski-Harabasz* memiliki hasil yang lebih baik dibandingkan *Dunn*. Hasil tersebut didapat dari algoritma *clustering* K-Means, K-Means Enhanced dan K-Means MMCA sebagai pembanding hasil validitas *cluster* (Khairatia et al., 2019).

4. Tidak memiliki fasilitas buang air besa/bersama-sama rumah tangga lain menggunakan satu jamban
5. Sumber penerangan rumah tangga tidak menggunakan listrik
6. Air minum berasal dari sumur/mata air yang tidak terlindung/sungai/air hujan
7. Bahan bakar untuk memasak sehari-hari adalah kayu bakar/arang/minyak tanah
8. Hanya mengonsumsi daging/susu/ayam satu kali dalam seminggu
9. Hanya membeli satu stel pakaian baru dalam setahun
10. Hanya mampu makan satu/dua kali dalam sehari
11. Tidak sanggup membayar biaya pengobatan di puskesmas/politeknik
12. Sumber penghasilan kepala rumah tangga adalah: petani dengan luas lahan 0,5 ha, buruh tani, nelayan, buruh bangunan, buruh perkebunan, atau pekerjaan lainnyadengan pendapatan di bawah Rp 600.000 per bulan
13. Pendidikan terakhir kepala rumah tangga: tidak sekolah/ tidak tamat sekolah dasar (SD)/hanya SD
14. Tidak memiliki tabungan/barang yang mudah dijual dengan nilai minimal Rp 500.000 seperti sepeda motor (kredit/nonkredit), emas, hewan ternak, kapal motor ataupun barang modal lainnya.

2.2. Kemiskinan dalam Perspektif Islam

Dalam Al-Quran kata miskin juga dikatakan sebagai “Maskanat” (Kemiskinan) yang merupakan *mashdar* dari *fi'il madhi* “Sakana”. Menurut Al-Raghib

2.3. Analisis *Clustering*

Clustering merupakan pengelompokan data/vektor secara partisi berdasarkan karakteristik masing-masing. Data-data hanya masuk pada *cluster* yang sama jika memiliki karakter yang sama. *Clustering* juga biasa disebut dengan *unsupervised learning* karena tidak memiliki ketentuan kelas (Jiawei et al., 2011).

Kelas pada *clustering* tidak memerlukan label ketika pemrosesan, karena label bisa diberikan setelah pengelompokan terbentuk. Oleh sebab itu, *clustering* cocok digunakan pada data yang sulit dicari label kelasnya ketika pembangkitan fitur. Selain itu, *clustering* juga biasa digunakan pada data yang akan dilakukan *supervised* untuk mengarakteristikan suatu data. Perlu diketahui bahwa, pendefinisian kemiripan (*similarity*) dalam *cluster* harus didasarkan terhadap atribut objek. Algoritma *clustering* yang digunakan tergantung pada proses yang diterapkan pada set data (Prasetyo, 2014).

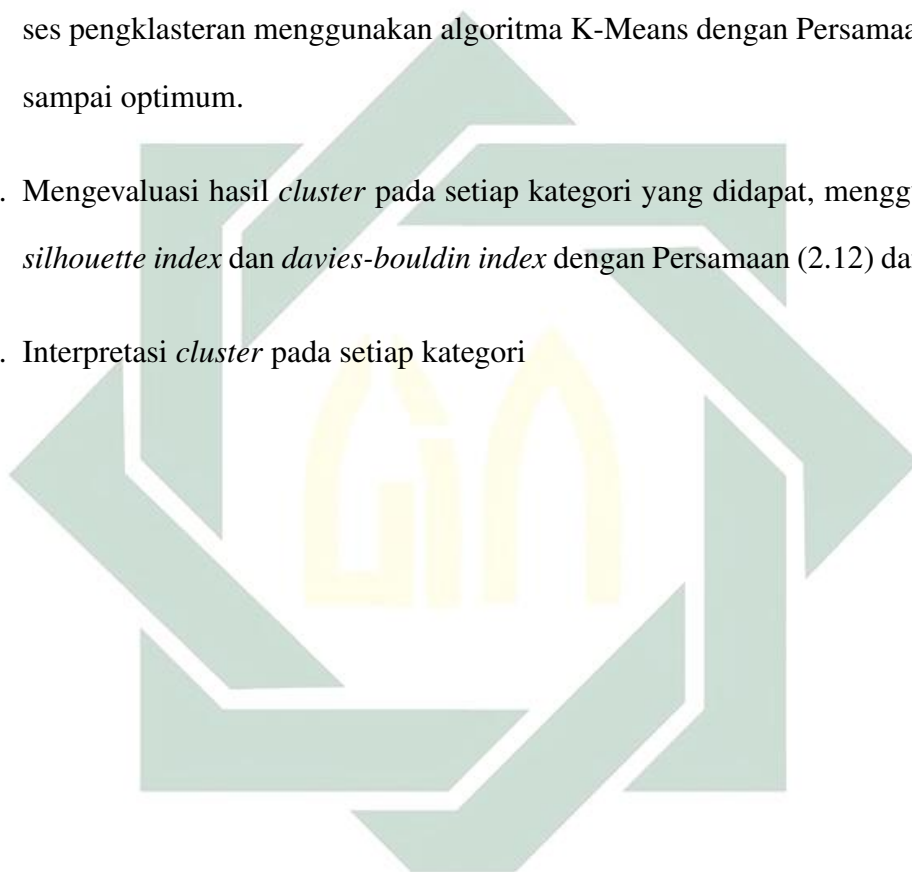
2.4. K-Means

K-Means merupakan salah satu metode *clustering* non-hierarki yang berfungsi untuk membagi data secara partisi. Pengelompokan data dalam metode K-Means tidak memerlukan kelas, sehingga data-data dikelompokkan langsung sesuai dengan masukan. Adapun masukan yang dimaksud yaitu data dan jumlah K *cluster* yang diinginkan. Data-data yang telah dimasukkan akan dikelompokkan ke dalam K *cluster* yang telah ditentukan. Di dalam setiap *cluster* terdapat *centroid* (titik pusat) yang dipilih secara acak pada tahap pertama. *Centroid* tersebut digunakan untuk mempresentasikan atau menggambarkan *cluster*. Posisi *centroid* dalam *cluster* akan terus berubah-ubah sampai pada posisi optimal (Wanto dkk., 2020). Pada setiap K *cluster* digambarkan oleh titik tunggal dalam r-dimensi (R^d). Set yang

			3 : Kolam/Sawah/Sungai/Danau/Laut 4 : Lubang Tanah 5 : Pantai/Tanah Lapang/Kebun 6 : Lainnya
Pangan	x5	Melewatkan Makan	1 : Iya 5 : Tidak 8 : Tidak Tahu 9 : Menolak Menjawab
	x6	Makan Lebih Sedikit	1 : Iya 5 : Tidak 8 : Tidak Tahu 9 : Menolak Menjawab
	x7	Kehabisan Makanan	1 : Iya 5 : Tidak 8 : Tidak Tahu 9 : Menolak Menjawab
Kesehatan	x8	Memiliki BPJS PBI	1 : Iya 5 : Tidak
	x9	Memiliki BPJS Non PBI	1 : Iya 5 : Tidak
	x10	Memiliki JAMKESMAS	1 : Iya 5 : Tidak
	x11	Memiliki Asuransi Kesehatan Swasta	1 : Iya 5 : Tidak
	x12	Memiliki Asuransi	1 : Iya

		Kesehatan Perusahaan	5 : Tidak
	x13	Tidak Punya Asuransi Kesehatan	1 : Iya 5 : Tidak
Karakteristik Kepala Rumah Tangga	x14	Lapangan Pekerjaan	0 : Tidak memiliki lapangan pekerjaan 1 : Pertanian Tanaman pada Palawija 2 : Hortikultura 3 : Perkebunan 4 : Perikanan ... 26 : Aktivitas Internasional
	x15	Pendidikan Terakhir Kepala Rumah Tangga	1 : Tidak Punya Ijazah SD 2 : Paket A 3 : SDLB 4 : SD 5 : MI 6 : SDLB 7 : SMPLB ... 21 : S3
Kepemilikan Aset	x16	Memiliki Kulkas	1 : Iya 5 : Tidak
	x17	Memiliki Ac	1 : Iya 5 : Tidak
	x18	Memiliki Mobil	1 : Iya 5 : Tidak

1. Memasukkan data pada setiap kategori yang telah ditentukan.
2. Menentukan jumlah K *cluster* pada setiap kategori yang akan digunakan, menggunakan metode *elbow* dengan Persamaan (2.3).
3. Menentukan *centroid* pada setiap kategori secara acak, untuk dilakukan proses pengklasteran menggunakan algoritma K-Means dengan Persamaan (2.1) sampai optimum.
4. Mengevaluasi hasil *cluster* pada setiap kategori yang didapat, menggunakan *silhouette index* dan *davies-bouldin index* dengan Persamaan (2.12) dan (2.4).
5. Interpretasi *cluster* pada setiap kategori



Tabel 4.6 Data Kepemilikan Aset

Data Ke-i	x16	x17	x18
1	1	5	5
2	1	5	5
3	5	5	5
4	1	1	5
5	5	5	5
6	5	5	1
7	1	1	1
8	1	1	1
9	1	1	1
10	1	1	5

Tabel 4.7 Centroid Pada Iterasi ke-1

Centroid	x16	x17	x18
c1	1	5	5
c2	5	5	5
c3	5	5	1
c4	1	1	1

Setelah menentukan *centroid* secara acak, kemudian menghitung jarak setiap data terhadap *centroid* menggunakan *Euclidean* pada Persamaan 2.1. Berikut merupakan contoh perhitungan jarak data ke-1 terhadap setiap *centroid*:

$$\begin{aligned}d(1,1) &= \sqrt{\sum_{i=1}^r (x_i, y_i)^2} \\&= \sqrt{(1-1)^2 + (5-5)^2 + (5-5)^2} \\&= \sqrt{0} \\&= 0\end{aligned}$$

$$\begin{aligned}d(1,2) &= \sqrt{\sum_{i=1}^r (x_i, y_i)^2} \\&= \sqrt{(1-5)^2 + (5-5)^2 + (5-5)^2} \\&= \sqrt{16} \\&= 4\end{aligned}$$

$$\begin{aligned}d(1,3) &= \sqrt{\sum_{i=1}^r (x_i, y_i)^2} \\&= \sqrt{(1-5)^2 + (5-5)^2 + (5-1)^2} \\&= \sqrt{32} \\&= 5.656854249\end{aligned}$$

$$\begin{aligned}d(1,4) &= \sqrt{\sum_{i=1}^r (x_i, y_i)^2} \\&= \sqrt{(1-1)^2 + (5-1)^2 + (5-1)^2} \\&= \sqrt{32} \\&= 5.656854249\end{aligned}$$

Tabel 4.8 Hasil Perhitungan Jarak Pada Iterasi ke-1

Data ke-i	Jarak ke centroid				Cluster yang diikuti
	c1	c2	c3	c4	
1	0	4	5.656854249	5.656854249	1
2	0	4	5.656854249	5.656854249	1
3	4	0	4	6.92820323	2
4	4	5.656854249	6.92820323	4	1
5	4	0	4	6.92820323	2
6	5.656854249	4	0	5.656854249	3
7	5.656854249	6.92820323	5.656854249	0	4
8	5.656854249	6.92820323	5.656854249	0	4
9	5.656854249	6.92820323	5.656854249	0	4
10	4	5.656854249	6.92820323	4	1

Setelah menghitung jarak setiap data, kemudian menentukan rata-rata setiap data yang bergabung pada *cluster* yang diikuti. Perhitungan rata-rata ini bertujuan untuk menentukan, apakah *centroid* pada iterasi ke-1 mengalami perubahan atau tidak. Berikut merupakan perolehan rata-rata pada setiap *cluster*:

Tabel 4.9 Rata-Rata *Cluster* ke-1 Pada Iterasi ke-1

Data k-i	x16	x17	x18
1	1	5	5
2	1	5	5
4	1	1	5
10	1	1	5
Rata-rata	1	3	5

Tabel 4.10 Rata-Rata *Cluster* ke-2 Pada Iterasi ke-1

Data ke-i	x16	x17	x18
3	5	5	5
5	5	5	5
Rata-rata	5	5	5

Tabel 4.11 Rata-Rata *Cluster* ke-3 Pada Iterasi ke-1

Data kw-i	x16	x17	x18
6	5	5	1
Rata-rata	5	5	1

Tabel 4.12 Rata-Rata *Cluster* ke-4 Pada Iterasi ke-1

Data ke-i	x16	x17	x18
7	1	1	1
8	1	1	1
9	1	1	1
Rata-rata	1	1	1

Dilihat dari nilai rata-rata pada setiap *cluster* yang diperoleh, terdapat perubahan *centroid* pada *cluster* ke-1. Untuk memastikan apakah terdapat objek yang berpindah atau tidak pada setiap *cluster*, maka diperlukan pengklasteran pada iterasi ke-2 menggunakan *centroid* baru yang ditunjukkan pada Tabel 4.13.

Tabel 4.13 *Centroid* Pada Iterasi ke-1

Centroid	x16	x17	x18
c1	1	3	5
c2	5	5	5
c3	5	5	1
c4	1	1	1

dengan cara sama, pengklasteran yang dilakukan pada iterasi ke-2 menghasilkan nilai sebagaimana pada Tabel 4.14.

Tabel 4.14 Hasil Perhitungan Jarak Pada Iterasi ke-2

Data ke-i	Jarak ke centroid				Cluster yang diikuti
	c1	c2	c3	c4	
1	2	4	5.656854249	5.656854249	1
2	2	4	5.656854249	5.656854249	1
3	4.472135955	0	4	6.92820323	2
4	2	5.656854249	6.92820323	4	1
5	4.472135955	0	4	6.92820323	2
6	6	4	0	5.656854249	3
7	4.472135955	6.92820323	5.656854249	0	4
8	4.472135955	6.92820323	5.656854249	0	4
9	4.472135955	6.92820323	5.656854249	0	4
10	2	5.656854249	6.92820323	4	1

Pada iterasi ke-2, objek pada setiap *cluster* tidak mengalami perubahan. Oleh sebab itu, proses iterasi berhenti dengan hasil akhir *centroid* yang terdapat pada Tabel 4.13. Pada *centroid* dapat diketahui rumah tangga pada kategori sangat

Tabel 4.21 Hasil *Cluster* Pada Kategori Pangan

Kategori	Cluster ke-i	Jumlah Rumah Tangga
Miskin	3	38
Cukup	1	37
Kaya	2	995

4.5.3. Kategori Kesehatan

Pengelompokan rumah tangga miskin berdasarkan pada hasil K-Means *cluster*, dengan pemilihan K *cluster* menggunakan metode *elbow* didapat hasil *centroid* pada tabel 4.22. Pada kategori ini, kesehatan suatu rumah tangga dilihat melalui asuransi kesehatan yang dimiliki. Adapun nilai variabel 0 yang didapat pada suatu *cluster* menunjukkan, bahwasannya rumah tangga yang berada *cluster* tersebut tidak memiliki asuransi kesehatan tersebut. Adapun keterangan setiap variabel dapat dilihat pada Tabel 3.1.

Tabel 4.22 Hasil *Centroid* Pada Kategori Kesehatan

Centroid	Jumlah	x8	x9	x10	x11	x12	x13
Cluster ke-i	Rumah Tangga						
1	195	1.8	4.30	4.16	4.53	4.96	5
2	101	4.76	1	4.92	4.64	4.72	1
3	86	4.63	5	4.91	4.86	1	4.02
4	290	5	1	4.88	4.81	4.67	5
5	313	5	5	4.88	4.80	5	1
6	85	1	5	4.95	5	4.39	1

Pada Tabel 4.22 dapat diketahui karakteristik rumah tangga berdasarkan ka-

Tabel 4.26 Hasil Centroid Pada Kategori Kepemilikan Aset

Centroid Cluster ke-i	Jumlah Rumah Tangga	x16	x17	x18
1	198	1	1.80	1
2	519	1	5	5
3	259	5	4.86	4.94
4	94	1	1	5

Pada Tabel 4.26 dapat diketahui karakteristik rumah tangga berdasarkan kategori kepemilikan aset. Adapun karakteristik rumah tangga yang tergolong pada *cluster 1* rata-rata memiliki kulkas, mayoritas memiliki AC dan memiliki mobil. Pada *cluster 2* rata-rata rumah tangga memiliki kulkas, tidak memiliki AC dan tidak memiliki mobil. Rumah tangga pada *cluster 3* tidak memiliki kulkas, mayoritas tidak memiliki AC dan mayoritas tidak memiliki mobil. Pada *cluster 4* rata-rata rumah tangga memiliki kulkas memiliki AC dan tidak memiliki mobil. Berdasarkan karakteristik tersebut dapat diketahui tingkat kemiskinan berdasarkan kategori kepemilikan aset sebagai berikut:

Tabel 4.27 Hasil Centroid Pada Kategori Kepemilikan Aset

Kategori	Cluster ke-i	Jumlah Rumah Tangga
Sangat Miskin	3	259
Miskin	2	519
Kaya	4	94
Sangat Kaya	1	198

nilai *davies-bouldin index* sebesar 0.33 dan *silhouette index* sebesar 0.96. Pada kategori kesehatan, diperoleh nilai *davies-bouldin index* sebesar 0.65 dan *silhouette index* sebesar 0.65. Pada kategori karakteristik rumah tangga, diperoleh nilai *davies-bouldin index* sebesar 0.61 dan *silhouette index* sebesar 0.53. Pada kategori kepemilikan aset, diperoleh nilai *davies-bouldin index* sebesar 0.29 dan *silhouette index* sebesar 0.92. Hasil validasi berdasarkan *silhouette index* yang lebih dari 0.5 dan *davies-bouldin index* yang mendekati 0 pada setiap kategori menunjukkan *cluster* yang terbetuk dalam penelitian ini layak digunakan.

5.2. Saran

Saran untuk penelitian lebih lanjut, dapat dilakukan penambahan variabel seperti jumlah penghasilan dan jumlah pengeluaran rumah tangga dalam satu bulan untuk menghasilkan pengelompokan rumah tangga miskin yang lebih akurat. Untuk memperoleh hasil yang lebih optimal dapat menggunakan metode *clustering* lainnya seperti K-Means++ dan K-Medoid.

- Prasetyo E., 2014, *Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab*. Yogyakarta: ANDI.
- Wanto A., Muhammad, N.H.S., Agus, P.W., Dedy, H., Ni Luh, W., Sri, R.G., Darmawan, N., Edi, S.N, Muhammad, R.L., Sarini, V.D., dan Cahyo, P., 2020, *Data Mining Algoritma dan Impementasi*, 1st ed., United States of America.
- Tan P., Michael, S., and Vipin, K., 2006. *Introduction to Data Mining*, 1st ed., United States of America.
- Purnama, B., 2019, *Pengantar Mechine Learning*. Bandung: Informatika.
- Fhadli, M., 2020, *Data Mining Dengan Python Untuk Pemula*. Ternate: Guepedia.
- Harani, N.H., Cahyo, dan P., Fikri, A.N., 2020, *Segmentasi Pelanggan Produk Digital Service Indihome Menggunakan AI-Goritma K-Means Berbasis Python*, *Jurnal Manajemen Informatika*.
- Bisri, A., 2020, *Terjemah Alfaraidul Bahiyyah*, Rembang: Menara Kudus.