ANALISIS GARIS KEMISKINAN MAKANAN MENGGUNAKAN METODE ALGORITMA K-MEANS CLUSTERING

Kamila Aprilia[1)], Falentino Sembiring[2)]
[1,2)]Program Studi Sistem Informasi, Universitas Nusa Putra
Jl. Raya Cibatu Cisaat No.21, Cibolang Kaler, Kec. Cisaat, Kab. Sukabumi, Jawa Barat 43155
e-mail : kamila.aprilia_si18@nusaputra.ac.id[1)], falentino.sembiring@nusaputra.ac.id[2)]

*Korespondensi: e-mail: falentino.sembiring@nusaputra.ac.id

ABSTRAK

Kemiskinan merupakan suatu permasalahan global yang banyak terjadi di berbagai negara salah satunya adalah negara Indo- nesia. Salah satu ciri adanya kondisi kemiskinan adalah ketidakmampuan masyarakat untuk memenuhi kebutuhan hidup seperti sandang, pangan dan papan. Kurangnya lahan produktif untuk menghasilkan aset pendapatan menjadi salah satu hal yang banyak ditemui di dalam permasalahannya terutama dalam memperoleh kebutuhan dasar salah satunya adalah makanan. Garis kemiskinan makanan merupakan nilai minimum dari pengeluaran kebutuhan makanan. Di dalam nya terdapat paket komoditi yang terdiri dari 52 jenis bahan makanan. Hal ini merupakan kebutuhan yang sangat penting bagi masyarakat. Hal ini juga menjadi salah satu faktor pemicu permasalahan kemiskinan di Indonesia dimana angka kematian setiap tahunnya meningkat dikarenakan kondisi kurangnya kebutuhan makanan di lingkungan masyarakat. Pada penelitian ini data yang digunakan ber- sumber dari Badan Pusat Statistik. Tujuan dari penelitian ini adalah untuk mengetahui tinggi, sedang dan rendahnya jumlah garis kemiskinan makanan berdasarkan provinsi di Indonesia dengan

etode Algoritma K-means Clustering. Berdasarkan hasil pengujian yang dilakukan dengan menerapkan Algoritma K-Means Clustering didapatkan hasil dengan Clus- ter 1 berjumlah 6 provinsi dengan kriteria garis kemiskinan makanan tertinggi, Cluster 2 berjumlah 16 provinsi dengan kriteria garis kemiskinan makanan sedang dan Cluster 3 berjumlah 12 provinsi dengan kriteria garis kemiskinan terendah. Hasil yang di dapat dari penelitian ini diharapkan mampu membantu pemerintah dalam mengambil keputusan, mendapatkan perhatian lebih terhadap cluster yang ditentukan.

***Kata Kunci**: Data Mining, Garis Kemiskinan Makanan, K-Means Clustering.*

ABSTRACT

Poverty is a global problem that occurs in many countries, one of which is Indonesia. One of the characteristics of the existence of poverty conditions is the inability of the community to meet the necessities of life such as clothing, food and shelter. The lack of productive land to generate income assets is one of the many problems encountered in the problem, especially in obtaining basic needs, one of which is food. The food poverty line is the minimum value of expenditure on food needs. Inside is a commodity package consisting of 52 types of food ingredients. This is a very important need for the community. This is also one of the triggering factors for the problem of poverty in Indonesia where the death rate increases every year due to the lack of food needs in the community. In this study, the data used were sourced from the Central Statistics Agency. The purpose of this study was to determine the high, medium and low number of food poverty lines based on provinces in Indonesia using the K-means Clustering Algorithm method. Based on the results of tests carried out by applying the K-Means Clustering Algorithm, the results obtained with Cluster 1 totaling 6 provinces with the highest food poverty line criteria, Cluster 2 totaling 16 provinces with moderate food poverty line criteria and Cluster 3 totaling 12 provinces with poverty line criteria Lowest. The results obtained from this study are expected to be able to assist the government in making decisions, getting more attention to the specified cluster.

***Keywords**: Data Mining, Food Proverty Line, K-Means Clustering*

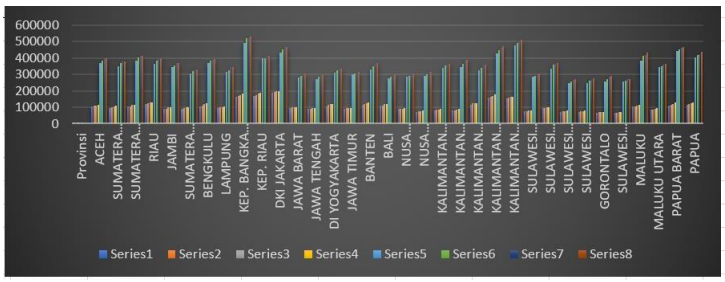
I. PENDAHULUAN

Garis kemiskinan atau batas kemiskinan adalah tingkat minimum pendapatan yang dianggap perlu dipenuhi untuk memperoleh standar hidup yang mencukupi di suatu negara. Garis kemiskinan

berguna sebagai perangkat ekonomi yang dapat digunakan untuk mengukur rakyat miskin dan mempertimbangkan pembaharuan sosio-ekonomi, misalnya seperti program peningkatan kesejahteraan dan asuransi pengangguran untuk menanggulangi kemiskinan [1]. Kemiskinan merupakan salah satu masalah yang terjadi di berbagai Negara termasuk di negara Indonesia. Ketidakmampuan masyarakat dalam memenuhi kebutuhan sandang, pangan dan papan membuat angka kemiskinan di Indonesia semakin meningkat.

Yang ditemukan di dalam permasalahannya terutama dalam memperoleh kebutuhan dasar salah satunya adalah makanan. Garis kemiskinan pada makanan merupakan nilai pengeluaran kebutuhan yang minimum pada makanan. Garis kemiskinan pada makanan juga disetarakan dengan 2100 kilo kalori perkapita perhari. Di dalamnya terdapat paket komoditi yang terdiri dari 52 jenis seperti padi-padian, umbi-umbian, ikan, daging, telur dan susu, sayuran, kacang-kacangan, buah-buahan, minyak dan lemak [2]. Hal ini merupakan salah satu kebutuhan pangan yang sangat penting bagi masyarakat di Indonesia. Disisi lain hal ini juga yang menjadi salah satu pemicu permasalahan yang paling utama pada kondisi kemiskinan dimana rata-rata masyarakat Indonesia masih banyak yang kelaparan dan kekurangan makanan sehingga menimbulkan angka kematian semakin meningkat setiap tahunnya. Namun demikian, masalah ini dapat diatasi dengan lebih ditingkatkan kembali sumber daya pada manusia agar masyarakat memiliki peluang lebih besar untuk meningkatkan setiap kebutuhan dan bertambahnya kesejahteraan hidup di dalamnya.

Pada penelitian ini, penulis akan membuat penelitian dalam mengelompokkan garis kemiskinan makanan berdasarkan provinsi di Indonesia menggunakan penerapan data mining dengan metode algoritma K-means Clustering. Adapun data yang digunakan dalam penelitian ini bersumber dari Badan Pusat Statistik. Hasil data yang diperoleh dan yang akan diolah adalah data garis kemiskinan makanan di Indonesia berdasarkan provinsi dari tahun 2017 – 2018.



Gambar 1. Grafik Garis Kemiskinan Makanan Berdasarkan Provinsi di Indonesia

II. TINJAUAN PUSTAKA

A. *Data Mining*

Data Mining merupakan teknologi baru yang sangat berguna untuk membantu perusahaan menemukan informasi yang sangat penting dari gudang data mereka. Beberapa aplikasi data mining fokus pada pred- iksi, mereka meramalkan apa yang akan terjadi dalam situasi baru dari data yang menggambarkan apa yang terjadi di masa lalu [3].

Data mining juga merupakan suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode data mining ini dapat digunakan untuk mengambil keputusan di masa depan. Data mining adalah pengolahan data dengan skala besar, sehingga data mining memiliki peranan penting dalam bidang industri, keuangan, cuaca, ilmu dan teknologi [4].

B. *Algoritma K-Means Clustering*

a) K-Means

Data Mining merupakan teknologi baru yang sangat berguna untuk membantu perusahaan menemukan informasi yang sangat penting dari gudang data mereka. Beberapa aplikasi data mining fokus pada prediksi, mereka meramalkan apa yang akan terjadi dalam situasi baru dari data yang menggambarkan apa yang terjadi di masa lalu [5].

Dalam penyelesaiannya, algoritma K-Means akan menghasilkan titik centroid yang di-jad- ikan tujuan dari algoritma K-Means. Setelah iterasi K-Means berhenti, setiap objek dalam da- taset menjadi anggota dari suatu cluster. Nilai cluster ditentukan dengan mencari seluruh objek untuk menemukan cluster dengan jarak terdekat ke objek . Algoritma K –means akan menge- lompokan item data dalam suatu dataset ke suatu cluster berdasarkan jarak terdekat [6].

b) Clustering

Clustering adalah proses pengelompokan benda serupa ke dalam kelompok yang berbeda, atau lebih tepatnya partisi dari sebuah data set kedalam subset, sehingga data dalam setiap subset memiliki arti yang bermanfaat. Dimana sebuah cluster terdiri dari kumpulan benda-benda yang mirip antara satu dengan yang lainnya dan berbeda dengan benda yang terdapat pada cluster lainnya. Algoritma clustering terdiri dari dua bagian yaitu secara hirarkis dan secara partitional. Algoritma hirarkis menemukan cluster secara berurutan dimana cluster ditetapkan sebelumnya, sedangkan algoritma partitional menentukan semua kelompok pada waktu tertentu [5].

C. Penelitian Terkait

Hasil penelitian yang digunakan penulis sebagai tinjauan pustaka di dalam penelitian ini adalah penelitian tentang data mining yang menggunakan penerapan metode K-means clustering yaitu oleh Disty Wahyuli, Handrizal, Iin Parlina, Agus Perdana Windarto, Dedi Suhendro, Anjar Wanto (2019) STIKOM Tunas Bangsa Pematangsiantar menulis jurnal yang berjudul " Mengelompokkan Garis Kemiskinan Menurut Provinsi Menggunakan Algoritma K-Medoids " menyatakan bahwa pada penelitian ini data yang digunakan bersumber dari badan pusat statistik. Tujuan dari penelitian ini adalah untuk mengetahui tinggi rendahnya jumlah kasus kemiskinan berdasarkan provinsi menggunakan k-medoids. Di Indonesia tingkat kemiskinan tinggi terdiri dari 23 provinsi dan tingkat kemiskinan rendah terdiri dari 11 provinsi. Hasilnya menyatakan bahwa Penerapan Data Mining menggunakan algoritma K-Medoids untuk mengelompokkan garis kemiski- nan menurut provinsi dengan data uji sebanyak 34 provinsi menghasilkan dua cluster yaitu cluster 1 (nilai terendah) sebanyak 11 dan cluster 2 (nilai tinggi) sebanyak 23. Penulis mengharapkan bahwa hasil dari penelitian ini dapat memberikan masukan kepada pemerintah dalam meningkatkan lapangan pekerjaan, sehingga dapat memperbaiki perekonomian masyarakat di Indonesia [7].

Penelitian selanjutnya adalah jurnal yang berjudul " Analisis Clustering Provinsi di Indonesia Berdasar- kan Tingkat Kemiskinan Menggunakan Algoritma K-Means" yang disusun dan ditulis oleh Achmad Ba- hauddin, Agustina Fatmawati, Febrianti Permata Sari (2021) Jurusan Teknik Industri, Fakultas Teknik, Universitas Sultan Ageng Tirtayasa menyatakan bahwa penerapan data mining dengan menggunakan algo- ritma K-means dapat diterapkan. Sumber data yang digunakan adalah Data kemiskinan yang diperoleh dari Badan Pusat Statistik. Pada penelitian ini digunakan metode clustering dengan algoritma K-Means untuk pengelompokan provinsi berdasarkan tingkat kemiskinannya dengan bantuan software Weka. Hasil penelitian menunjukkan terdapat 3 cluster provinsi di Indonesia berdasarkan tingkat kemiskinannya yaitu Cluster 0 (provinsi dengan tingkat kemiskinan rendah), Cluster 1 (provinsi tingkat kemiskinan sedang), dan Cluster 2 (provinsi dengan tingkat kemiskinan tinggi). Provinsi yang termasuk dalam kategori provinsi dengan tingkat kemiskinan tinggi yaitu Maluku, Papua Barat dan Papua [8].

Penelitian lain yang menjadi tinjauan pustaka di dalam penulisan penelitian ini yaitu oleh Suhartini, Ria Yuliani (2021) Program Studi Teknik Informatika Universitas Hamzanwadi menulis jurnal berjudul "Pen- erapan Data Mining Untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma K-Means di Dusun Bagik Endep Sukamulia Timur". Sumber data dari penelitian ini diperoleh dari hasil ob- servasi/pengamatan ke kantor desa Sukamulia Timur, wawancara yang dilakukan dengan staf Kantor Desa Sukamulia Timur untuk mengetahui gambaran tentang kondisi perekonomian Penduduk dan studi pustaka untuk menunjang metode wawancara dan observasi yang telah dilakukan. Simpulan data yang digunakan adalah data penduduk Sukamulia Timur pada tahun 2019 yang berjumlah 200 data dengan 9 atribut yaitu nama penduduk, pekerjaan, penghasilan/bulan, jumlah anak yang sekolah SD, jumlah anak yang sekolah SMP, jumlah anak yang sekolah SMA,

jumlah anak yang kuliah, jumlah anak yang tidak sekolah serta jumlah anggota keluarga. Berdasarkan hasil pengujian yang dilakukan dengan menerapkan algoritma K- Means didapatkan hasil dengan Cluster 1 berjumlah 18 penduduk dengan kriteria Penduduk ekonomi tinggi, Cluster 2 berjumlah 72 Penduduk dengan kriteria Penduduk ekonomi sedang, dan Cluster 3 berjumlah 110 penduduk dengan kriteria Penduduk ekonomi rendah [9].

Penelitian lain yang terakhir dijadikan tinjauan pustaka di dalam penulisan penelitian ini yaitu oleh Falentino Sembiring, Octaviana, Sudin Saepudin (2020) Universitas Nusa Putra menulis jurnal yang ber- judul "Implementasi Metode K-Means dalam Pengklusteran Daerah Pungutan Liar di Sukabumi (Studi Kasus : Dinas Kependudukan dan Pencatatan Sipil)". Sumber data yang diambil dalam penelitian ini adalah

data laporan pada bulan Januari 2019 berbentuk SQL Penelitian ini hanya akan membahas tentang berapa banyak kejadian pungli yang terjadi di setiap kecamatan yang ada di kabupaten Sukabumi tentang kependudukan dan pencatatan sipil. Adapun tujuan dari penelitian ini adalah untuk menentukan cluster tingkat pungutan liar tinggi, sedang dan rendah. Hasil dari penelitian ini memperoleh data indeks kecamatan dengan tingkat jumlah laporan masyarakat daerah terhadap pungutan liar, data dengan klasifikasi tingkat tinggi yaitu Cireunghas, Gegerbitung, Kalapa Nunggal, Kalibunder, Purabaya, Simpenan, Parung Kuda, Sukaraja, Nagrak, Nyalindung, Pelabuhanratu, Surade, Warungkiara. Data tingkat pungutan liar sedang terdapat pada 19 kecamatan dan 15 tingkat pungutan liar rendah. [10]

III. METODOLOGI PENELITIAN

Tujuan dari penelitian ini adalah untuk mengelompokkan garis kemiskinan makanan dari tahun 2017-2020 berdasarkan provinsi di Indonesia dengan menggunakan metode K-Means Clustering. Beberapa langkah yang harus di lakukan adalah sebagai berikut :

A. Analisis Data

Proses ini dapat dilakukan ketika peneliti sudah mengumpulkan data. Data pada penelitian ini di peroleh dari data yang dikumpulkan berdasarkan dokumen tentang kemiskinan dan ketimpangan pada Badan Pusat Statistik melalui situs https://www.bps.go.id. Data ini merupakan data sekunder yaitu data garis kemiskinan makanan berdasarkan kalkulasi dari pedesaan dan perkotaan dari 34 provinsi di Indonesia.

Tabel 1. Data Garis Kemiskinan Makanan

No	Provinsi	2017		2018		2019		2020	
1	ACEH	104496	107572	111335	114830,4	370093	384381	397032	398316
2	SUMATERA UTARA	97102	101729	104865	107932,34	351215	367105	376790	378617
3	SUMATERA BARAT	106715	107368	112319	116426	382333	402003	413832	414949
4	RIAU	120571	122833	128099	131734,43	365515	381904	396883	398654
5	JAMBI	88019	90001	100071	100501	343783	355585	368137	369835
6	SUMATERA SELATAN	90775	93568	99621	101928,35	306087	318123	327021	328710
7	BENGKULU	106779	107735	117725	125225,38	369211	384886	389613	392261
8	LAMPUNG	95176	98583	101073	102954,58	313620	326364	342151	344944
9	KEP. BANGKA BELITUNG	162685	166347	173633	182017,46	492693	521298	524371	528771
10	KEP. RIAU	166969	173687	181902	188369,72	396007	400070	410225	410811
11	DKI JAKARTA	189163	195054	198949	198987,31	429915	451918	466156	467847

12	JAWA BARAT	95922	99879	101223	102311,77	281693	292718	301806	305687
...	.....	.....	.....	.....	.....	.....	.....	.....	.....
...	.....	.....	.....	.....	.....	.....	.....	.....	.....
32	MALUKU UTARA	84356	86985	88628	95381,02	346075	348361	358264	362406
33	PAPUA BARAT	110377	114383	117927	128283,36	441803	453232	463545	465715
34	PAPUA	112904	117951	124711	129967,53	404631	418367	423243	437370

B. Pengolahan Data

Adapun langkah-langkah proses perhitungan menggunakan metode algoritma K-means clustering adalah sebagai berikut :

a. Tentukan jumlah cluster K. Adapun jumlah cluster yang di buat pada penelitian ini terdiri dari 3 cluster yaitu tertinggi, sedang dan terendah.
b. Pilih pusat cluster centroid awal (Iterasi 1). Biasanya untuk menentukan pusat cluster dilakukan secara acak (random). Pada penelitian ini penentuan titik pusat cluster ditentukan dari yang tertinggi ke yang terendah.
c. Jarak antara data dan pusat cluster di hitung dengan menggunakan rumus Euclidean Distance.

$$d(x_j, c_j) = \sqrt{\sum_{j=1}^{n} (x_j - c_j)^2}$$

Dimana :

d : Jarak

Xj : Data ke j

Cj : Centroid ke j

Tabel 2. Titik Pusat Awal Tiap Cluster

Centroid 1	162685	166347	173633	182017	492693	521298	524371	528771
Centroid 2	94388	97377	100595	101178,63	336739	357664	363287	370310
Centroid 3	63493	66374	69333	72578	254518	262966	270655	271458

Tabel 3. Hasil Perhitungan Jarak Terdekat Cluster pada Iterasi 1

No	C1	C2	C3	Jarak Terdekat	Cluster
1	286786,1	65247,04	259019,8	65247,04	2
2	326821,2	25301,69	218712	25301,69	2
3	257405,7	95989,1	289959,3	95989,10	2
4	277592	79821,01	268286	79821,01	2
5	350377,9	13167,82	196498,3	13167,82	2
6	421228,1	74685,28	124174,1	74685,28	2
7	289071,9	64053,73	255854	64053,73	2
8	397150	51086,21	148309,2	51086,21	2

9	0,46	351349,5	544732,7	0,46	1
10	226206,2	185630,8	354990,9	185630,83	2
11	135312,4	274634,7	456494,6	135312,41	1
12	465341,7	123346,6	88549,23	88549,23	3
……	……	……	……	……	……
……	……	……	……	……	……
15	365839,2	25433,57	182686,1	25433,57	2
16	162251,3	202250,8	395204,5	162251,29	1
17	216456,8	136244,1	330083,6	136244,11	2

Tabel 4. Hasil Cluster Iterasi 1

Cluster	Provinsi	Hasil
1	9,11,23,24,33	5
2	1,2,3,4,5,6,7,8,10,14,16,20,21,22,26,31,32,34	18
3	12,13,15,17,18,19,25,27,28,29,30	11

d. Setelah semua titik ditempatkan di cluster terdekat, maka hitung kembali ke titik pusat cluster berdasarkan perhitungan rata-rata anggota yang ada pada cluster tersebut dengan menggunakan rumus :

Keterangan :

$$C = \frac{\Sigma m}{n}$$

C : Centroid data
m : anggota data yang masuk ke dalam centroid tertentu
n : jumlah data yang menjadi anggota pada centroid tertentu

Tabel 5. Titik Pusat Cluster pada Iterasi 2

Centroid 1	155017	159984,8	164471,4	170779,38	454371,2	472451,6	483389	488324,2
Centroid 2	103755,3	107335,1	112047,3	115775,62	351443,7	365732	376960,2	380752,7
Centroid 3	81367,82	83791,64	86347,82	88574,629	272680	281758,5	292415,6	296032,5

Tabel 6. Hasil Perhitungan Jarak Terdekat Cluster pada Iterasi 2

No	C1	C2	C3	Jarak Terdekat	Cluster
1	204136	37535,76	209383,1	37535,76263	2
2	243697,9	13975,88	169345,5	13975,88262	2
3	176494,7	69339,08	240861,1	69339,07697	2
4	192438,5	47034,05	216238,5	47034,04526	2
5	267737,2	35826,55	148660,2	35826,55239	2
6	336478	101145,4	72403,87	72403,86733	3
7	205570,9	33255,43	205444,8	33255,43113	2
8	312456,8	76872,85	96771,54	76872,84516	2
9	86484,13	320823	493491,1	86484,12702	1
10	144743,2	154023,2	301261,8	144743,184	1

11	78206,24	242058,3	402805,3	78206,24441	1
12	379794,4	147968,3	35543,48	35543,47741	3
……	……	……	……	……	……
……	……	……	……	……	……...
15	283544,1	52645,88	136509,4	52645,88478	2
16	97349,16	175559,4	347036,8	97349,16413	1
17	135646,4	107261,9	280199,1	107261,9327	2

Tabel 7. Hasil Cluster pada Iterasi 2

Cluster	Provinsi	Hasil
1	9,10,11,23,24,33	6
2	1,2,3,4,5,7,8,14,16,20,21,22,26,31,32,34	16
3	6,12,13,15,17,18,19,25,27,28,29,30	12

e. Proses K-Means akan terus beriterasi dan dilakukan secara berulang-ulang, dengan kata lain proses perhitungan terus dilakukan hingga menghasilkan pengelompokkan data yang sama dan tidak berubah dengan hasil iterasi data sebelumnya. Pada penelitian ini perhitungan iterasi di lakukan sebanyak 3 kali, proses perhitungannya berhenti pada iterasi ke 3 yang dapat di lihat hasilnya pada tabel di bawah ini.

Tabel 8. Titik Pusat Cluster pada Iterasi 3

Centroid 1	157009	162268,5	167376,5	173711,1	444643,8	460388	471195	475405,3
Centroid 2	100615,7	104048,6	108458,1	112103,9	351493,3	366561,4	378002,4	382126,8
Centroid 3	82151,75	84606,33	87453,92	89687,44	275463,9	284788,9	295299,4	298755,7

Tabel 9. Hasil Perhitungan Jarak Terdekat Cluster pada Iterasi ke 3

No	C1	C2	C3	Jarak Terdekat	Cluster
1	187442,6	36481,33	203381,9	36481,33364	2
2	226272,8	7886,688	163356,8	7886,688376	2
3	161534,5	68191,38	234873,7	68191,37713	2
4	173268	50831,32	210230,2	50831,32472	2
5	250588,5	31453,53	142747,4	31453,5311	2
6	317164,4	101236,2	66370,21	66370,21172	3
7	187809,4	34595,41	199416,6	34595,40825	2
8	293400,7	76914,29	90770,49	76914,28781	2
9	108863,8	321991,4	487467,6	108863,7788	1
10	120619,3	159462	295588,6	120619,32	1
11	64153,55	245794	396987,2	64153,54532	1
12	359529,4	148867,4	31046,52	31046,5194	3
……	……	……	……	……	……
……	……	……	……	…….	……
15	266591,7	48657,53	130719,4	48657,52821	2
16	95855,11	174716	341065	95855,10592	1
17	121474,3	107394,2	274187,7	107394,1971	2

Tabel 10. Hasil Perhitungan Akhir Cluster pada Iterasi 3

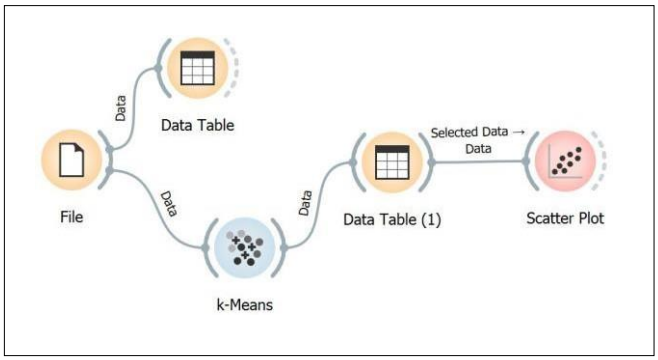
Cluster	Provinsi	Hasil
1	9,10,11,23,24,33	6
2	1,2,3,4,5,7,8,14,16,20,21,22,26,31,32,34	16
3	6,12,13,15,17,18,19,25,27,28,29,30	12

Pada Iterasi ke 3 ini titik pusat dari seriap cluster sudah tidak berubah dan data tidak ada lagi yang berpin- dah dari satu cluster ke cluster lainnya.

IV. HASIL DAN PEMBAHASAN

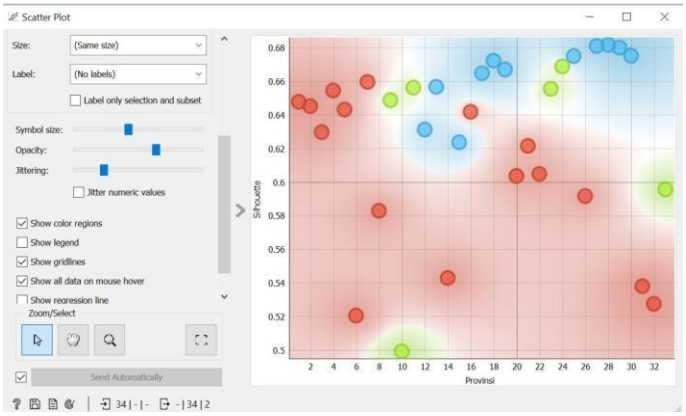
A. Hasil Iterasi K-Means dengan software Orange

Pengolahan data garis kemiskinan makanan berdasarkan provinsi di Indonesia dengan menggunakan soft- ware Orange bisa di lihat pada gambar di bawah ini :



Gambar 1. Pemodelan Pengolahan Data K-Means Menggunakan Software Orange

Dengan menggunakan pemodelan seperti gambar di atas, dengan 34 data dan 3 cluster sesuai dengan pendefinisian K maka di peroleh cluster 1 sebanyak 6 data, cluster 2 sebanyak 16 data dan cluster 3 sebanyak 12 data. Berikut adalah gambar Scatter plot dari setiap cluster.



Gambar 2. Model Scatter Plot pada Software Orange

Pada gambar di atas grafik scatter plot terdapat 2 variabel yaitu X dan Y dimana X adalah Provinsi dan Y adalah Silhouette. Di dalam scatter plot terdapat pengelompokkan jumlah bulatan yang menyebar dari 34 data yang dibedakan oleh 3 warna yaitu warna hijau muda, merah dan biru muda. Untuk warna hijau muda yaitu cluster 1 dengan jumlah 6 bulatan yang menandakan bahwa di

dalam cluster 1 adalah cluster provinsi garis kemiskinan makanan tertinggi dengan jumlah 6 data, Warna merah menandakan cluster 2 sebagai cluster provinsi sedang untuk garis kemiskinan makanan dengan jumlah data 16. Dan yang terakhir adalah bulatan warna biru muda yang memiliki 12 bulatan menandakan bahwa bulatan ini adalah 12 data pada cluster provinsi terendah untuk garis kemiskinan makanan berdasarkan provinsi di Indonesia.

V. KESIMPULAN

Berdasarkan pembahasan dan hasil implementasi menggunakan software Orange pada data garis kem- iskinan makanan berdasarkan provinsi di Indonesia, maka dapat diambil kesimpulan :

a. Penelitian data garis kemiskinan makanan berdasarkan provinsi di Indonesia yang diolah dalam Microsoft Excel dengan menggunakan metode K-means clustering dan diimplementasikan pada software Orange menghasilkan nilai centroid pada 3 cluster yaitu tertinggi, sedang dan terendah.
b. Hasil yang diperoleh dari penelitian dengan menggunakan metode K-means clustering ini menghasilkan 6 provinsi untuk cluster garis kemiskinan makanan tertinggi yaitu Kep. Bangka Be- litung, Kep. Riau, DKI Jakarta, Kalimantan Timur, Kalimantan Utara dan Papua Barat, 16 provinsi untuk cluster garis kemiskinan makanan sedang yaitu Aceh, Sumatera Utara, Sumatera Utara, Riau, Jambi, Bengkulu, Lampung, DI Yogyakarta, Banten, Kalimantan Barat, Kalimantan Tengah, Kali- mantan Selatan, Sulawesi Tengah, Maluku, Maluku Utara dan Papua. Sedangkan untuk cluster garis kemiskinan makanan rendah sebanyak 12 provinsi yaitu Sumatera Selatan, Jawa Barat, Jawa Ten- gah, Jawa Timur, Bali, Nusa Tenggara Timur, Nusa Tenggara Barat, Sulawesi Utara, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat.
c. Hasil yang di dapat dari penelitian ini diharapkan mampu membantu pemerintah dalam mengambil keputusan dan menemukan informasi yang dibutuhkan untuk memecahkan masalah dalam mendata garis kemiskinan makanan berdasarkan provinsi serta dapat mengetahui provinsi mana saja yang harus mendapatkan perhatian lebih terhadap cluster yang ditentukan.

DAFTAR PUSTAKA

[1] D. RI, "Garis Kemiskinan," DPR RI, [Online]. Available: https://berkas.dpr.go.id. [Accessed 19 Juni 2021].

[2] Badan Pusat Statistik, "Garis Kemiskinan Makanan," BPS, [Online]. Available: http://bps.go.id. [Accessed 19 Juni 2021].

[3] I. H. Witten, E. Frank and M. A. Hall, "Data Mining Practial Machine Learning Tools And Techniques," 2011.

[4] J. O. Ong, "Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University," *Jurnal Ilmiah Teknik Industri*, no. 12, pp. 10-20, 2013.

[5] T. S. Madhulata, "Data Structures and Algorithms," *IOSR Journal of Engineering*, vol. 2, no. 4, pp. 719-725, 2012.

[6] B. M. N. Bangoria and V. Pambhar, "A Survey on Efficient Enchanted K-Means Clustering Algorithm," *International Journal For Scientific Research and Development*, vol. I, no. 9, pp. 1698- 700, 2013.

[7] D. Wahyuli , Handrizal, l. Parlina, A. P. Windarto, D. Suhendro and A. Wanto, "Mengelompokkan Garis Kemiskinan Menurut Provinsi Menggunakan Algoritma K-Medoids," *Prosiding Seminar Nasional Riset Information (SENARIS)*, pp. 452-461, 2019.

[8] A. Bahauddin, A. Fatmawati and F. P. Sari, "ANALISIS CLUSTERING PROVINSI DI INDONESIA BERDASARKAN TINGKAT KEMISKINAN MENGGUNAKAN ALGORITMA K-MEANS," *MISI (Jurnal Manajemen Informatika dan Sistem Informasi)*, vol. 4, no. 1, 2021.

[9] S. and R. Yuliani, *Infotek (Jurnal Informatika dan Teknologi)*, vol. 4, no. 1, pp. 39-50, 2021.

[10] F. Sembiring, O. and S. Saepudin, "IMPLEMENTASI METODE K-MEANS DALAM PENGKLASTERAN DAERAH PUNGUTAN LIAR DI KABUPATEN SUKABUMI (STUDI KASUS : DINAS KEPENDUDUKAN DAN PENCATATAN SIPIL)," *Jurnal Tekno Insentif*, vol. 14, no. 1, pp. 40-47, 2020.