

PENERAPAN METODE K-MEANS DALAM PENGKLASTERAN WILAYAH DI INDONESIA BERDASARKAN PEMBERIAN ASI EKSKLUSIF PADA BAYI

¹Zulia Imami Alfianti*, ²Ginabila, ³Ahmad Fauzi

^{1,2,3}Fakultas Teknik & Informatika Universitas Bina Sarana Informatika

^{1,2,3}Jl. Kramat Raya No.98, Jakarta Pusat, DKI Jakarta 10450

¹zulia.zim@bsi.ac.id, ²gina.gnb@bsi.ac.id, ³ahmad.aau@bsi.ac.id

*) Penulis Korespondensi

Abstrak

Air Susu Ibu (ASI) adalah cairan yang keluar dari kelenjar payudara seseorang ibu yang memiliki beragam gizi yang diperlukan dalam menopang perkembangan serta pertumbuhan balita. Air Susu Ibu (ASI) Eksklusif merupakan pemberian ASI yang tidak disertai dengan pemberian suplementasi makanan maupun minuman lain kecuali obat. Saat ini pemberian asi eksklusif dipengaruhi banyak faktor yaitu ibu yang bekerja, pendidikan ibu yang rendah, gencarnya periklan tentang penggunaan susu formula, asi yang tidak keluar dan banyak faktor lainnya menyebabkan tidak semua bayi mendapatkan asi eksklusif. Pengklasteran wilayah dilakukan berdasarkan persentase data pemberian asi eksklusif pada bayi usia 6 bulan dari 35 provinsi di Indonesia. Pengklasteran dilakukan untuk mengelompokkan 35 provinsi di Indonesia menjadi wilayah provinsi kasus tinggi, sedang, dan rendah. Hasil dari penelitian ini yaitu 31% provinsi memiliki persentase tertinggi, 40% memiliki persentase sedang dan 29 % memiliki persentase rendah.

Kata Kunci: asi eksklusif, klasifikasi, k-means

Abstract

Breast milk is a fluid that comes out of a mother's breast glands which contains a variety of nutrients needed to support the development and growth of toddlers. Exclusive breast milk (ASI) is breastfeeding that is not accompanied by any other food or drink supplementation except medication. Currently, exclusive breastfeeding is influenced by many factors, namely working mothers, low maternal education, incessant advertising about the use of formula milk, breast milk not coming in and many other factors causing not all babies to receive exclusive breast milk. In this research, regional clustering will be carried out based on the percentage of exclusive breastfeeding for 6 month old babies from 34 provinces in Indonesia. Clustering was carried out to group 34 provinces in Indonesia into provinces with high, medium and low cases. The results of this research are that 31% of provinces have the highest percentage, 40% have a medium percentage and 29% have a low percentage.

Keywords: breast milk, clasification, k-means

PENDAHULUAN

Air Susu Ibu (ASI) adalah cairan yang berasal dari kelenjar payudara seseorang ibu yang menjadi asupan utama bayi yang memiliki beragam gizi yang diperlukan dalam

menopang perkembangan serta pertumbuhan bayi [1]. Air Susu Ibu (ASI) Eksklusif merupakan pemberian ASI sebagai asupan utama yang tidak disertai dengan pemberian suplementasi makanan maupun minuman lain kecuali obat [2]. Pemberian ASI Eksklusif

dilakukan sejak bayi baru dilahirkan sampai dengan 6 bulan pertama, setelah masa tersebut ASI tidak mampu memenuhi kebutuhan mineral dan gizi sehingga harus disertai dengan Makanan Pendamping Air Susu Ibu (MP ASI) [3]. Saat ini pemberian asi eksklusif dipengaruhi banyak faktor yaitu ibu yang bekerja, pendidikan ibu yang rendah, maraknya iklan tentang penggunaan susu formula, asi yang tidak keluar dan banyak faktor lainnya menyebabkan tidak semua bayi mendapatkan ASI eksklusif [4].

Bayi yang tidak mendapatkan ASI secara eksklusif sampai usia 6 bulan cenderung mengalami pertumbuhan kurang optimal yang rentan terindikasi stunting dan kurangnya ikatan emosional ibu dengan bayi [5]. Dampak dari rendahnya pemberian ASI eksklusif pada bayi 0-6 bulan yaitu dapat memperberat penyakit seperti ISPA 35.09%, diare 38.07%, dan gizi kurang 49,2% yang dapat menimbulkan beberapa efek negatif pada bayi seperti lambatnya pertumbuhan badan, rawan terhadap penyakit, menurunnya tingkat kecerdasan dan terganggunya mental anak, kekurangan gizi yang serius dapat menyebabkan kematian anak [6]. Pemerintah khususnya Dinas Kesehatan Republik Indonesia tentu diharapkan dapat memberikan penyuluhan optimal kepada ibu dan calon ibu untuk mengetahui pentingnya pemberian ASI eksklusif pada bayi sampai usia 6 bulan. Kualitas ASI juga sangat perlu diperhatikan dalam pemberian ASI eksklusif mengingat kualitas ASI juga mempengaruhi gizi dan

vitamin yang akan diterima oleh bayi melalui ASI sebagai modal dasar pertumbuhan manusia di awal kehidupan [7].

Data mining adalah suatu proses pengambilan data, penggunaan data yang pernah digunakan untuk mendapatkan suatu susunan yang sistematis dari data dalam jumlah besar [8]. Hasil pengolahan data tersebut umumnya berfungsi sebagai alat pengambil keputusan. Proses data mining dalam pengelompokan beberapa data ke dalam sebuah kelompok disebut Teknik *clustering* [9]. *Clustering* digunakan untuk melakukan pengelompokan data provinsi yang memiliki kasus stunting ke dalam sejumlah kelompok [10]. Penelitian ini bertujuan untuk menerapkan metode *K-Means* dalam pengklasteran wilayah di Indonesia berdasarkan pemberian ASI eksklusif pada bayi dimana metode tersebut akan mengelompokkan wilayah-wilayah di Indonesia ke dalam beberapa klaster berdasarkan tingkat keberhasilan pemberian ASI eksklusif. Penelitian ini mengidentifikasi pola distribusi wilayah dengan tingkat pemberian ASI eksklusif yang tinggi, sedang, atau rendah sehingga hasil penelitian diharapkan dapat memberikan informasi yang relevan bagi pemerintah atau lembaga kesehatan dalam merancang kebijakan atau program yang lebih terfokus untuk meningkatkan cakupan ASI eksklusif, terutama di wilayah-wilayah yang masih memiliki tingkat pemberian ASI eksklusif yang rendah. Berdasarkan tujuan penelitian

yang telah diuraikan, timbul beberapa pertanyaan terkait bagaimana metode *K-Means* dapat diterapkan untuk mengklusterisasi wilayah di Indonesia berdasarkan tingkat pemberian ASI eksklusif pada bayi, dan apa saja pola yang dapat diidentifikasi dari hasil pengklasteran tersebut. Penelitian ini juga berupaya menjawab pertanyaan terkait jumlah kluster optimal yang dapat menggambarkan variasi tingkat pemberian ASI eksklusif, serta karakteristik setiap kluster yang terbentuk.

Studi literatur telah dilakukan berdasarkan penelitian sebelumnya yaitu yang dilakukan oleh Nainggolan dan Hendikawati pada tahun 2024, penelitian ini mengelompokkan Kabupaten atau Kota di Provinsi Jawa Tengah berdasarkan faktor-faktor yang mempengaruhi stunting dengan metode *average linkage*. Hasil dari penelitian ini mengetahui bahwa faktor-faktor pembeda utama antar cluster adalah Vitamin A dan ASI eksklusif, yang memiliki pengaruh signifikan [11]. Penelitian kedua dilakukan oleh Candra, Yanuar dan Devianto pada tahun 2023.

Data yang digunakan adalah data masalah gizi balita di Indonesia pada tahun 2022. Penelitian ini berhasil mengelompokkan kabupaten/kota di Indonesia berdasarkan masalah gizi balita menggunakan metode *Two Step Cluster* yang menghasilkan 2 *cluster*, dan metode *Ensemble K-Modes* yang menghasilkan 5 *cluster* dengan karakteristik berbeda [12].

Penelitian ketiga dilakukan oleh Wang Dkk pada tahun 2020 dengan tujuan untuk memprediksi keberhasilan pemberian ASI eksklusif (*exclusive breastfeeding, EBF*) hingga dua bulan setelah melahirkan menggunakan *decision tree* (DT). Model DT menunjukkan akurasi 73,1%, sensitivitas 75,5%, dan spesifisitas 66,3%, yang berarti cukup andal dalam memprediksi risiko berhentinya EBF sebelum dua bulan [13]. Penelitian selanjutnya dilakukan oleh Seegard Dkk pada tahun 2024. Penelitian ini berisi tentang hubungan antara durasi pemberian ASI eksklusif pada bayi dan risiko kanker pada anak dengan menggunakan data dari 309.473 anak yang lahir di Denmark antara 2005 hingga 2018. Penelitian ini mendukung bukti sebelumnya bahwa ASI eksklusif lebih lama dapat membantu melindungi anak dari kanker tertentu, dan menekankan perlunya penelitian lebih lanjut untuk memahami mekanisme biologis di balik hubungan tersebut [14]. Penelitian kelima dilakukan oleh Roig Dkk pada tahun 2024, dimana penelitian ini bertujuan untuk memprediksi pemberian ASI eksklusif selama masa rawat inap ibu setelah melahirkan menggunakan model *Machine Learning* (ML) dan teknik *Explainable Artificial Intelligence* (XAI) untuk menjelaskan hasil prediksi tersebut. Teknik XAI digunakan untuk menjelaskan hubungan antara faktor-faktor tersebut dengan hasil prediksi serta menunjukkan variasi hasil berdasarkan karakteristik individu [15].

METODE PENELITIAN

Tahapan penelitian yang dilakukan oleh peneliti adalah mulai dari identifikasi tujuan penelitian, Pengumpulan Data, Perbaikan Data, Pemilihan dan Penerapan Algoritma sampai dengan Evaluasi Hasil Penelitian. Gambar 1 adalah kerangka tahapan penelitian yang dilakukan oleh peneliti:

1. Identifikasi Tujuan Penelitian

Pada tahap ini peneliti mulai mendefinisikan tujuan penelitian, yaitu untuk mengelompokkan wilayah di Indonesia berdasarkan tingkat pemberian ASI eksklusif pada bayi. Tujuan utamanya adalah untuk mengetahui pola-pola pengelompokan wilayah yang memiliki tingkat pemberian ASI eksklusif yang tinggi, sedang, atau rendah.

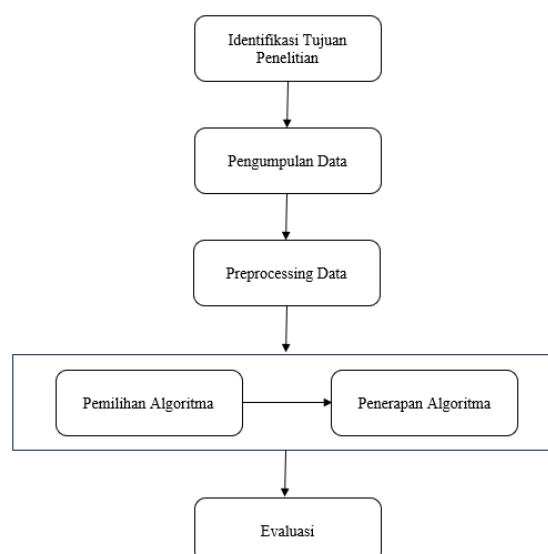
2. Pengumpulan Data

Data yang digunakan dalam penelitian merupakan data sekunder yang diperoleh

dari data sensus atau laporan kesehatan dari Badan Pusat Statistik (BPS) atau Kementerian Kesehatan Republik Indonesia. Data yang dikumpulkan mencakup berbagai variabel yang mempengaruhi pemberian ASI eksklusif yang berjumlah 35 data provinsi di Indonesia, seperti tingkat pendidikan ibu, akses ke fasilitas kesehatan, kondisi sosial ekonomi, prevalensi ASI eksklusif di setiap daerah, dan faktor-faktor lain yang relevan.

3. Preprocessing Data

Setelah data terkumpul, dilakukan tahap preprocessing untuk memastikan kualitas data. Tahap ini meliputi pembersihan data (menghapus nilai yang hilang atau data yang tidak *relevan*), normalisasi atau standarisasi data (untuk menghindari dominasi atribut tertentu dalam analisis), serta penanganan data kategorikal.



Gambar 1. Kerangka Penelitian

4. Pemilihan Algoritma

Pada tahap ini metode *K-Means* digunakan untuk menentukan jumlah kluster. Pada tahap ini peneliti menentukan jumlah kluster yang optimal yang dapat menggambarkan pola pemberian ASI eksklusif. *Elbow Method* atau *Silhouette Analysis* digunakan untuk memilih jumlah kluster yang tepat untuk penelitian.

5. Penerapan Algoritma

Metode *K-Means* akan melakukan iterasi untuk mengelompokkan data ke dalam kluster berdasarkan kedekatannya dengan centroid yang telah diinisialisasi. Pengelompokkan pola menjadi beberapa *cluster* menggunakan rumus *euclidean distance*. Dengan menghitung jarak antara data dengan *centroid* sama dengan akar dari hasil penjumlahan antara data pada atribut ke-*n* dikurangi objek data pada *centroid* ke-*k*. Hasil jarak ini yang akan digunakan untuk mencari data mana yang memiliki tingkat kemiripan berdasarkan jarak yang paling kecil/terpendek. Perhitungan diformulasikan dengan menggunakan persamaan *euclidean distance* (1) yang menghasilkan jarak data dengan titik pusat *cluster* :

$$d(a, b) = \sqrt{\sum_{k=1}^n (a_k - b_k)^2} \quad (1)$$

Dalam persamaan (1), jarak data dengan titik pusat *cluster* diperoleh dari kuadrat jumlah dimensi (atribut) atribut ke- *k* dari objek data.

6. Evaluasi

Setelah pengklasteran selesai, dilakukan tahap evaluasi untuk menganalisis hasil kluster dan memahami karakteristik masing-masing kluster. Peneliti akan melihat apakah kluster-kluster yang terbentuk menggambarkan pola yang relevan dalam pemberian ASI eksklusif di setiap wilayah.

HASIL DAN PEMBAHASAN

Data awal yang digunakan dalam penelitian ini sebanyak 35 provinsi yang ada di Indonesia yang kemudian dibagi kedalam tiga kluster untuk mengetahui tingkat pemberian asi eksklusif yang tinggi, sedang, dan rendah. Tabel 1 menunjukkan data tingkat pemberian ASI eksklusif di setiap provinsi di Indonesia dari tahun 2021 hingga 2023. Berdasarkan data awal yang terdapat pada Tabel 1 maka dipilih tiga pusat kluster secara acak berdasarkan provinsi yang memiliki persentase pemberian asi tinggi, sedang, dan menengah yaitu K1, K18, K29 sebagai pusat kluster awal untuk pengolahan data pada iterasi awal. Data centroid awal tersebut dapat dilihat dalam Tabel 2.

Tabel 1. Data Bayi Usia Kurang dari 6 Bulan yang Mendapatkan ASI Eksklusif Menurut Provinsi

Kode	Provinsi	2021	2022	2023
K1	Aceh	66.66	65.88	67.05
K2	Sumatera Utara	57.83	57.17	61.98
K3	Sumatera Barat	74.16	74.32	75.84
K4	Riau	70.29	69.51	71.14
K5	Jambi	71.37	72.68	74.14
K6	Sumatera Selatan	69.93	70.46	75.59
K7	Bengkulu	67.08	67.84	72.44
K8	Lampung	74.93	76.76	76.20
K9	Kep. Bangka Belitung	62.94	62.77	63.30
K10	Kep. Riau	58.84	62.40	62.88
...	...	...	...	...
K18	Nusa Tenggara Barat	81.46	79.69	82.45
K19	Nusa Tenggara Timur	81.18	78.56	78.74
K20	Kalimantan Barat	71.21	72.66	72.97
K21	Kalimantan Tengah	55.98	55.26	55.78
K22	Kalimantan Selatan	60.27	64.59	65.69
K23	Kalimantan Timur	75.87	76.58	77.70
K24	Kalimantan Utara	81.00	78.70	77.81
...	...	...	...	...
K29	Gorontalo	52.75	53.60	55.11
K30	Sulawesi Barat	74.75	74.67	75.04
K31	Maluku	61.32	59.62	61.52
K32	Maluku Utara	66.00	68.83	69.66
K33	Papua Barat	58.77	61.79	62.92
K34	Papua	74.08	74.18	55.41
K35	Indonesia	71.58	72.04	73.97

Sumber : bps.go.id

Tabel 2. Data Centroid Awal Iterasi 0

Centroid Awal	Atribut 1	Atribut 2	Atribut 3
C1	66.66	65.88	67.05
C2	81.46	79.69	82.45
C3	52.75	53.60	55.11

Nilai *centroid* awal pada Tabel 2. masing-masing *centroid* awal telah diperoleh dari hasil pemilihan seara acak dari 35 *dataset* dengan ketentuan nilai persentase terendah, sedang dan tertinggi. Penentuan *centroid* awal telah dilakukan, selanjutnya perhitungan jarak antara *centroid* awal dengan titik *centroid* menggunakan rumus sebagai berikut. Berdasarkan persamaan maka,

$$d_{(a,b)} = \sqrt{(66.66 - 66.66)^2 + (65.88 - 65.88)^2 + (67.05 - 67.05)^2}$$

$$d_{(a,b)} = \sqrt{(66.66 - 81.46)^2 + (65.88 - 79.69)^2 + (67.05 - 82.45)^2}$$

$$d_{(a,b)} = \sqrt{(66.66 - 52.75)^2 + (65.88 - 53.60)^2 + (67.05 - 55.11)^2}$$

Hasil yang didapatkan untuk menentukan jarak antara *centroid* awal dengan titik *centroid* menggunakan rumus *eucledian distance* yang hasilnya terdapat pada tabel 3 ini, kemudian dipilih *centroid* mana dari masing-masing kode yang memiliki jarak terpendek: Tabel 3. merupakan hasil

perhitungan menggunakan rumus *eucledian distance*, setelah itu dipilih *centroid* yang memiliki jarak terpendek. Tabel 3 merupakan hasil dari perhitungan yang dilakukan untuk pencarian jarak terpendek sehingga diperoleh anggota *cluster* yaitu C1 berjumlah 21 data, C2 berjumlah 11 data, dan C3 berjumlah 3 data, terdapat pada tabel 4.

Tabel. 3 Hasil Perhitungan Jarak *Cluster* dengan Titik *Centroid* Iterasi 0

Kode	C1	C2	C3
K1	0.00	25.43	22.06
K2	13.40	38.53	9.26
K3	14.31	11.22	36.30
K4	6.56	18.88	28.60
K5	10.89	14.83	32.76
K6	10.23	16.28	31.60
K7	5.75	21.15	26.62
K8	16.45	9.50	38.38
K9	6.13	31.56	15.97
K10	9.52	34.55	13.23
K11	9.49	21.04	28.36
K12	19.74	6.15	41.63
K13	22.09	3.52	44.11
K14	18.96	6.63	40.97
K15	7.43	18.31	29.20
K16	10.56	15.20	32.37
K17	2.77	22.84	24.67
K18	25.43	0.00	47.46
K19	22.54	3.89	44.61
K20	10.09	15.63	31.98
K21	18.81	44.24	3.69
K22	6.66	30.95	17.01
K23	17.68	7.97	39.66
K24	22.04	4.77	44.08
K25	6.75	31.85	15.72
K26	0.81	26.10	21.40
K27	17.28	8.21	39.31
K28	7.15	32.57	14.94
K29	22.06	47.46	0.00
K30	14.37	11.19	36.40
K31	9.91	35.31	12.28
K32	3.99	22.82	24.88
K33	9.80	34.88	12.82
K34	16.11	28.57	29.64
K35	10.49	15.10	32.41

Tabel 4. Keanggotaan Cluster Hasil Perhitungan pada Iterasi 0

Centroid		Keanggotaan Cluster																			
C1		K1	K4	K5	K6	K7	K9	K10	K11	K15	K16	K17	K20	K22	K25	K26	K31	K32	K33	K34	K35
C2		K3	K8	K12	K13	K14	K18	K19	K23	K24	K27	K30									
C3		K2	K21	K29																	

Tabel 4. menunjukkan hasil *cluster*, dan daftar Keanggotaan *cluster* keanggotaan klaster (cluster membership) menunjukkan provinsi-provinsi (K1, K2, K3, pada iterasi pertama (iterasi 0) dalam metode dll.). Langkah selanjutnya yaitu dengan *K-Means* yang digunakan untuk kembali menghitung nilai rata-rata anggota mengelompokkan data. Setiap *Centroid* (C1, *cluster* sehingga didapatkan nilai *centroid* C2, C3) mewakili pusat dari masing-masing iterasi 1 yang dapat dilihat pada tabel 5:

Tabel 5. Nilai *Centroid* pada Iterasi 1

Centroid Awal	Atribut 1	Atribut 2	Atribut 3
C1	66,42	67,19	68,36
C2	77,47	77,09	78,14
C3	55,52	55,34	57,62

Tabel 6. Hasil Perhitungan Jarak *Cluster* dengan Titik *Centroid* Iterasi 1

Kode	C1	C2	C3
K1	1.87	19.12	18.00
K2	14.66	32.31	5.26
K3	12.91	4.89	32.24
K4	5.30	12.57	24.53
K5	9.38	8.52	28.72
K6	8.68	10.36	27.55
K7	4.19	15.03	22.57
K8	15.02	3.21	34.36
K9	7.57	25.23	11.93
K10	10.51	28.21	9.40
K11	8.07	15.52	24.40
K12	18.29	2.19	37.57
K13	20.72	3.00	40.05
K14	17.58	0.49	36.92
K15	5.94	12.08	25.14
K16	9.06	9.01	28.31
K17	2.29	16.59	20.57
K18	24.10	6.42	43.38
K19	21.33	4.03	40.54
K20	8.61	9.25	27.96
K21	20.24	37.93	1.90
K22	7.19	24.64	13.16
K23	16.27	1.74	35.61
K24	20.84	3.89	40.03
K25	7.77	25.52	11.74
K26	2.19	19.76	17.35
K27	15.93	1.85	35.25
K28	8.62	26.27	10.87
K29	23.39	41.15	4.13
K30	13.04	4.78	32.35
K31	11.41	29.02	8.19
K32	2.13	16.48	20.90
K33	10.83	28.56	8.95
K34	16.59	23.16	26.54
K35	9.03	8.81	28.35

Nilai *centroid* pada tabel 5 dihasilkan dari proses iterasi 1 yang telah dilakukan menggunakan rumus *euclidian distance*. Nilai *centroid* pada tabel 5 ini kemudian diolah kembali dengan menghitung jarak *cluster* dengan titik pusat pada tabel 5 sehingga diperoleh hasil perhitungan jarak dengan titik pusat pada tabel 6. Setelah dilakukan perhitungan jarak *cluster* dengan titik pusat, diperoleh keanggotaan *cluster* pada iterasi 1 dengan jumlah masing-masing anggota pada tabel 7 yaitu C1 berjumlah 15, C2 berjumlah 15 dan C3 berjumlah 4 anggota *cluster*.

Perhitungan yang dilakukan pada iterasi 1 pada tabel 7 diperoleh keanggotaan *cluster* yang masih mengalami perubahan nilai, maka perhitungan masih harus dilakukan sampai keanggotaan *cluster* tidak mengalami perubahan (tetap) terhadap keanggotaan *cluster* sebelumnya. Perhitungan dengan rumus yang sama kembali dilakukan, maka

perhitungan terhenti di iterasi 7 karena keanggotaan *cluster* pada iterasi 7 sesuai dengan keanggotaan *cluster* pada iterasi sebelumnya (iterasi 6).

Berikut adalah hasil perhitungan iterasi 6 dan iterasi ke 7 : Nilai *centroid* pada iterasi 6 pada tabel 8 digunakan kembali untuk perhitungan jarak *cluster* dengan titik *centorid* yang dihasilkan poada tabel 9 sehingga diperoleh hasil perhitungan harak dengan titik *centroid* iterasi 6. Perhitungan jarak *cluster* dengan titik pusat diperoleh keanggotaan *cluster* pada iterasi 6 dengan jumlah masing-masing anggota pada tabel 10 yaitu C1 berjumlah 14, C2 berjumlah 11 dan C3 berjumlah 10 anggota *cluster*. Jumlah keanggotaan *cluster* iterasi 6 pada tabel 10 masih terjadi perubahan keanggotaan *cluster*, maka perhitungan iterasi masih kembali dilakukan dengan nilai *cenroid* pada iterasi 7 yang terdapat pada tabel 11.

Tabel 7. Keanggotaan Cluster Hasil Perhitungan pada Iterasi 1

Centro id	Keanggotaan Cluster														
	K	K4	K6	K7	K9	K1	K1	K1	K2	K2	K2	K2	K2	K3	K3
C1	1					1	5	7	0	2	5	6	8	2	4
C2	K3	K5	K8	K1	K1	K1	K1	K1	K1	K2	K2	K2	K2	K3	K3
	3			2	3	4	6	8	9	1	3	4	7	0	5
C3	K2	K1	K2	K3	K3										
	2	0	9	1	3										

Tabel 8. Nilai Centroid pada Iterasi 6

Centroid Awal	Atribut 1	Atribut 2	Atribuit 3
C1	68,63	68,97	69,91
C2	76,16	76,03	77,30
C3	59,23	60,20	61,66

Tabel 9. Hasil Perhitungan Jarak *Cluster* dengan Titik *Centroid* Iterasi 6

Kode	C1	C2	C3
K1	4.65	17.28	10.79
K2	17.85	30.44	3.36
K3	9.71	3.01	24.97
K4	2.13	10.72	17.29
K5	6.26	6.65	21.42
K6	6.01	8.53	20.34
K7	3.17	13.16	15.37
K8	11.83	1.81	27.06
K9	10.70	23.38	4.80
K10	13.73	26.34	2.54
K11	7.35	13.76	17.53
K12	15.14	2.96	30.30
K13	17.52	4.82	32.78
K14	14.38	1.69	29.64
K15	3.03	10.21	17.88
K16	6.01	7.13	21.04
K17	2.61	14.76	13.42
K18	20.90	8.24	36.14
K19	18.09	5.80	33.32
K20	5.44	7.39	20.65
K21	23.40	36.08	8.34
K22	10.34	22.77	6.05
K23	13.08	0.74	28.33
K24	17.61	5.55	32.82
K25	11.00	23.66	4.43
K26	5.23	17.92	10.10
K27	12.72	0.32	27.98
K28	11.75	24.42	3.86
K29	26.60	39.29	11.34
K30	9.81	3.00	25.08
K31	14.54	27.18	2.17
K32	2.65	14.61	13.57
K33	14.06	26.69	2.08
K34	16.35	22.07	21.33
K35	5.88	6.93	21.08

Tabel 10. Keanggotaan Cluster Hasil Perhitungan pada Iterasi 6

Centroid	Keanggotaan Cluster													
C1	K1	K4	K5	K6	K7	K11	K15	K16	K17	K20	K26	K32	K34	K35
C2	K3	K8	K12	K13	K14	K18	K19	K23	K24	K27	K30			
C3	K2	K9	K10	K21	K22	K25	K28	K29	K31	K33				

Tabel 11. Nilai Centroid Pada Iterasi 7

Centroid Awal	Atribut 1	Atribut 2	Atribuit 3
C1	69,22	69,65	70,84
C2	77,47	77,09	78,14
C3	59,23	60,20	61,66

Tabel 12. Hasil Perhitungan Jarak *Cluster* dengan Titik *Centroid* Iterasi 7

Kode	C1	C2	C3
K1	5.93	19.12	10.79
K2	19.08	32.31	3.36
K3	8.44	4.89	24.97
K4	1.12	12.57	17.29
K5	4.97	8.52	21.42
K6	4.87	10.36	20.34
K7	3.22	15.03	15.37
K8	10.58	3.21	27.06
K9	11.98	25.23	4.80
K10	14.95	28.21	2.54
K11	7.04	15.52	17.53
K12	13.85	2.19	30.30
K13	16.25	3.00	32.78
K14	13.11	0.49	29.64
K15	1.88	12.08	17.88
K16	4.72	9.01	21.04
K17	3.69	16.59	13.42
K18	19.63	6.42	36.14
K19	16.88	4.03	33.32
K20	4.19	9.25	20.65
K21	24.68	37.93	8.34
K22	11.50	24.64	6.05
K23	11.80	1.74	28.33
K24	16.41	3.89	32.82
K25	12.24	25.52	4.43
K26	6.52	19.76	10.10
K27	11.46	1.85	27.98
K28	13.03	26.27	3.86
K29	27.86	41.15	11.34
K30	8.57	4.78	25.08
K31	15.81	29.02	2.17
K32	3.52	16.48	13.57
K33	15.29	28.56	2.08
K34	16.80	23.16	21.33
K35	4.59	8.81	21.08

Tabel 13. Keanggotaan *Cluster* Hasil Perhitungan pada Iterasi 7

Centroid	Keanggotaan Cluster													
C1	K1	K4	K5	K6	K7	K11	K15	K16	K17	K20	K26	K32	K34	K35
C2	K3	K8	K12	K13	K14	K18	K19	K23	K24	K27	K30			
C3	K2	K9	K10	K21	K22	K25	K28	K29	K31	K33				

Perhitungan iterasi 7 yang telah dilakukan menggunakan rumus *euclidian distance* menghasilkan perhitungan jarak terpendek dengan titik *centroid* pada tabel 7. Hasil perhitungan iterasi 7 yang telah dilakukan menghasilkan keanggotaan *cluster* yang sama dengan keanggotaan *cluster* pada iterasi 6 yang terdapat pada tabel 13. Maka, perhitungan iterasi menggunakan rumus *euclidian distance* dihentikan pada iterasi 7

dengan total keanggotaan *cluster* pada tabel 13 dengan rincian sebanyak C1 14 anggota, C2 sebanyak 11 anggota, dan C3 sebanyak 10 anggota *cluster*.

Proses iterasi terhenti di iterasi 7 karena nilai *centroid* iterasi 7 sama dengan nilai *centroid* iterasi 6. Perbandingan nilai *centroid* yang berasal dari iterasi awal sampai dengan

iterasi 7 tersebut dapat dilihat pada tabel 10.

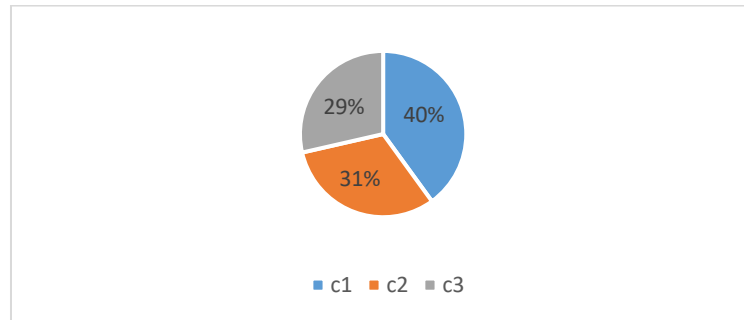
Perbandingan nilai *centroid* pada tabel 14 yang berisi nilai *centorid* pada masing-masing iterasi, maka proses iterasi berhenti pada iterasi 7 karena nilai *centroid* pada iterasi 7 sama dengan nilai *centroid* pada iterasi 6 sehingga didapatkan keanggotaan *cluster* seperti ditampilkan pada tabel 15:

Tabel 14. Perbandingan Nilai Centroid Iterasi Awal Sampai dengan Iterasi 7.

Centroid	Atribut 1	Atribut 2	Atribut 3
Iterasi Awal			
C1	66.66	65.88	67.05
C2	81.46	79.69	82.45
C3	52.75	53.60	55.11
Iterasi 1			
C1	66.42	67.19	68.35
C2	77.47	77.09	78.13
C3	55.52	55.34	57.62
Iterasi 2			
C1	66.78	67.39	68.36
C2	74.81	74.64	75.86
C3	57.90	58.91	60.88
Iterasi 3			
C1	67.41	68.01	68.62
C2	75.46	75.47	76.93
C3	59.12	59.72	61.21
Iterasi 4			
C1	68.20	68.39	68.94
C2	75.72	75.70	77.12
C3	59.23	60.20	61.66
Iterasi 5			
C1	68.63	68.97	69.91
C2	76.16	76.03	77.30
C3	59.23	60.20	61.66
Iterasi 6			
C1	69.22	69.65	70.84
C2	77.47	77.09	78.14
C3	59.23	60.20	61.86
Iterasi 7			
C1	69.22	69.65	70.84
C2	77.47	77.09	78.14
C3	59.23	60.20	61.86

Tabel 15. Keanggotaan Cluster Hasil Perhitungan pada Iterasi 7

Centroid	Keanggotaan Cluster													
C1	K1	K4	K5	K6	K7	K11	K15	K16	K17	K20	K26	K32	K34	K35
C2	K3	K8	K12	K13	K14	K18	K19	K23	K24	K27	K30			
C3	K2	K9	K10	K21	K22	K25	K28	K29	K31	K33				



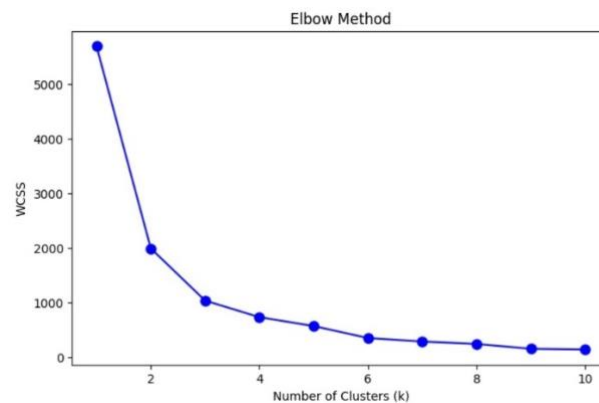
Gambar 2. Diagram Lingkaran Persentase Masing-Masing Cluster

Dari hasil *clustering* 35 provinsi di Indonesia berdasarkan penerima ASI eksklusif diatas, maka untuk lebih jelasnya penulis gambarkan pada sebuah diagram lingkaran yang ditampilkan pada gambar 1. Diagram lingkaran yang terdapat pada gambar 1 dijelaskan bahwa hasil klasterisasi yang dilakukan berdasarkan penentuan *cluster* awal yaitu sebanyak tiga *cluster* wilayah menghasilkan persentase perbandingan pada Gambar 2. Hasil perhitungan menggunakan metode *K-Means* pada diagram lingkaran yang terdapat pada Gambar 2, dari 35 provinsi di Indonesia yang telah melewati proses *clustering* menggunakan metode *K-Means* dibagi menjadi tiga *cluster* penerima asi eksklusif yang dikategorikan menjadi tiga provinsi tingkat penerima tinggi, provinsi yang memiliki tingkat penerima asi eksklusif sedang/menengah yang memiliki jumlah penerima diantara tinggi dan rendah, dan

provinsi yang memiliki tingkat penerima rendah. Diagram lingkaran diatas yang dihasilkan dari proses *clustering* terdapat 31% persentase C2 yaitu 11 provinsi yang memiliki tingkat penerima tinggi, 40% persentase *cluster* C1 terdiri dari 14 provinsi memiliki tingkat penerima menengah, dan 29% persentase *cluster* C3 terdiri dari 10 provinsi memiliki tingkat penerima rendah. Dari hasil Gambar 2. diatas, kemudian dilakukan metode optimasi menggunakan metode *elbow* untuk menentukan jumlah *cluster* terbaik. Berikut adalah hasil perhitungan SSE berdasarkan pengujian nilai $k=1$ sampai dengan $k=10$. Dari perhitungan SSE metode Elbow pada Tabel 4, didapatkan nilai SSE yang memiliki nilai terbesar terdapat pada nilai $k=2$, sehingga jumlah cluster optimal yang dapat dibentuk sejumlah 2 cluster. Perhitungan tersebut digambarkan dengan Grafik *Elbow Method* pada gambar 3.

Tabel 16. Nilai WCSS Untuk Berbagai Jumlah Cluster

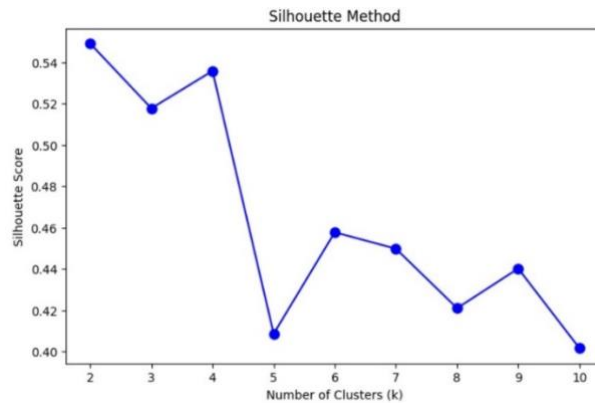
K	Nilai SSE
1	5695.67
2	1935.67
3	1029.36
4	725.37
5	484.81
6	322.62
7	232.56
8	181.65
9	155.60
10	130.79

**Gambar 3. Grafik Sse Elbow Method****Tabel 17. Nilai Silhouette Score Untuk Berbagai Jumlah Cluster**

K	Nilai SSE
2	0.545
3	0.518
4	0.536
5	0.517
6	0.467
7	0.477
8	0.485
9	0.409
10	0.406

Grafik Elbow menunjukkan bahwa jumlah cluster terbaik adalah ketika WCSS mulai menurun perlahan, biasanya di titik "siku". Pada pengelompokan data sebanyak 35 data menggunakan koefisien silhouette sebagai metode penentu dalam jumlah cluster optimal pada clustering Elbow Method menunjukkan hasil berupa jumlah cluster

optimal sebanyak 2 cluster. Dari perhitungan *Silhouette Score*, didapatkan nilai yang memiliki nilai terbesar terdapat pada nilai $k=2$, sehingga jumlah cluster optimal yang dapat dibentuk menurut metode *Silhouette Score* sejumlah 2 cluster. Perhitungan tersebut digambarkan dengan Grafik *Silhouette Score* pada gambar 3.



Gambar 4. Silhouette Method Untuk Berbagai Cluster

Nilai maksimum *Silhouette Score* adalah 0.545 untuk $K=2$, sehingga jumlah *cluster* terbaik menurut metode ini adalah 2. Terdapat beberapa aspek yang perlu diperhatikan sebagai Keterbatasan dalam penelitian mengenai penerapan metode *K-Means* dalam pengklasteran wilayah di Indonesia berdasarkan pemberian ASI eksklusif pada bayi. Penelitian ini sangat bergantung pada kualitas dan ketersediaan data, sehingga jika data yang digunakan tidak lengkap atau tidak representatif, hasil pengklasteran dapat tidak mencerminkan kondisi yang sebenarnya di lapangan. Pemilihan variabel juga menjadi faktor penting dalam penelitian ini, karena jika variabel yang dipilih tidak mencakup semua faktor yang memengaruhi pemberian ASI eksklusif, hasil pengklasteran bisa jadi tidak menggambarkan realitas secara menyeluruh. Keterbatasan lainnya adalah asumsi homogenitas dalam klaster, di mana *K-Means* menganggap setiap klaster memiliki karakteristik yang sama, padahal wilayah dalam satu klaster mungkin memiliki perbedaan signifikan dalam faktor sosial, ekonomi, dan budaya. Terakhir,

keterbatasan waktu dan sumber daya dalam penelitian ini juga dapat memengaruhi kedalaman analisis atau cakupan wilayah yang dianalisis, yang berpotensi membatasi kesimpulan yang dapat ditarik dari penelitian ini.

KESIMPULAN DAN SARAN

Proses pengelompokkan wilayah provinsi berdasarkan data bayi penerima ASI eksklusif di Indonesia dengan menggunakan metode *K-Means* dari data 35 provinsi yang ada di Indonesia 31% provinsi/wilayah termasuk kedalam kategori penerima asi eksklusif tinggi artinya tingkat penerimanya masih tergolong banyak, 40% wilayah termasuk kedalam kategori penerima ASI sedang, dan 29% wilayah termasuk kedalam kategori penerima ASI eksklusif rendah artinya wilayah tersebut yang harus lebih dulu menerima penanganan lanjutan dari dinas kesehatan republik Indonesia karena masih kurangnya pemahaman mengenai pentingnya memberikan asi eksklusif pada bayi. Mengingat pentingnya pemberian ASI

eksklusif pada bayi usia 0 sampai 6 bulan untuk menjadi tolak ukur kesehatan dan tumbuh kembang bayi, maka pemerintah harus memberikan penanganan tepat kepada ibu hamil pada daerah yang kurang pemahaman terhadap pentingnya pemberian ASI eksklusif seperti kegiatan penyuluhan dan pengarahan bagi para ibu hamil untuk mempersiapkan pemberian ASI eksklusif sejak awal kehamilan seperti makan makanan sehat dan didukung dengan pemberian vitamin yang seimbang.

Penelitian ini telah menghasilkan pola distribusi wilayah dengan tingkat pemberian ASI eksklusif yang tinggi, sedang, atau rendah sehingga hasil penelitian diharapkan dapat memberikan informasi yang relevan bagi pemerintah atau lembaga kesehatan dalam merancang kebijakan atau program yang lebih terfokus untuk meningkatkan cakupan ASI eksklusif, terutama di wilayah-wilayah yang masih memiliki tingkat pemberian ASI eksklusif yang rendah. Sebagai gerakan untuk menjadikan masyarakat yang sehat dan sumber daya manusia yang berkualitas di masa depan maka penelitian ini perlu dikembangkan dan diperbaharui setiap tahunnya berdasarkan data-data terbaru. Selain itu, dibutuhkan penelitian lanjutan metode *clustering* terhadap kasus penyakit atau kondisi yang umumnya menjadi permasalahan masyarakat Indonesia.

DAFTAR PUSTAKA

- [1] F. The, M. Hasan, and S. D. Saputra, "Edukasi Pentingnya Pemberian ASI Eksklusif pada Bayi di Puskesmas Gambesi," *J. Surya Masy.*, vol. 5, no. 2, p.208,2023,doi:10.26714/jsm.5.2.2023.208-213.
- [2] Zikrina, "Faktor-Faktor yang Mempengaruhi Dalam Pemberian ASI Eksklusif," *Idea Nurs. J.*, vol. XIII, no. 3, pp. 7–14, 2022, [Online]. Available: <https://jurnal.usk.ac.id/INJ/article/view/22690>
- [3] D. Parwati and K. U. Thalib, "Pentingnya Pemberian Asi Eksklusif Pada Bayi Di Desa Boda-Boda Kabupaten Mamuju," *Safari*, vol. 2, no. 3, pp. 123–127, 2022.
- [4] H. Febriyanti, "Faktor-Faktor Yang Berhubungan Dengan Pemberian Asi Eksklusif Pada Tenaga Kesehatan Yang Memiliki Bayi Di Wilayah Kabupaten Pringsewu Tahun 2017," *Midwifery J. J. Kebidanan UM. Mataram*, vol. 3, no. 1, p. 38, 2018, doi: 10.31764/mj.v3i1.125.
- [5] S. Sofiyati, "Penyuluhan ASI Eksklusif Dengan Kejadian Stunting Pada Usia Balita (1-5 Tahun) Di Desa Kalikoa Kecamatan Kedawung Kabupaten Cirebon," *Nat. J. Pelaks. Pengabd. Berger. bersama Masyarakat.*, vol. 2, no. 1, pp. 17–35, 2024.
- [6] F. J. Prihatini, K. Achyar, and I. R. Kusuma, "Faktor – Faktor yang Mempengaruhi Ketidakberhasilan ASI Eksklusif pada Ibu Menyusui," *J. Ris. Kesehat. Masy.*, vol. 3, no. 4, pp. 184–191,2023, doi:10.14710/jrkm.2023.18811

- [7] A. A. Yusuff, F. Fardhoni, E. L. Rehkliana, and R. Rahayu, "Studi Potong Lintang Pemberian ASI Eksklusif dan Faktor yang Mempengaruhinya," *J. Manaj. Kesehat. Yayasan RS.Dr. Soetomo*, vol. 8, no. 1, p. 178, 2022, doi: 10.29241/jmk.v8i1.954.
- [8] S. Handoko, F. Fauziah, and E. T. E. Handayani, "Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering," *J. Ilm. Teknol. dan Rekayasa*, vol. 25, no. 1, pp. 76–88, 2020, doi: 10.35760/tr.2020.v25i1.2677.
- [9] Z. I. Alfianti, "Pengelompokan Wilayah Kasus Balita Stunting Di Indonesia Menggunakan Algoritma K-Means," *J. Ilm. Inform. Komput.*, vol. 28, no. 3, pp. 173–185, 2023, doi: 10.35760/ik.2023.v28i3.8876.
- [10] Z. I. Alfianti, "Pengelompokan Wilayah Penyebaran Covid-19 Di Kabupaten Karawang Menggunakan Algoritma K-Means," *J. Ilm. Inform. Komput.*, vol. 26, no. 2, pp. 111–122, 2021, doi: 10.35760/ik.2021.v26i2.4155.
- [11] E. Novia and P. Hendikawati, "Average Linkage Hierarchical Cluster untuk Pengelompokan Kabupaten / Kota di Provinsi Jawa Tengah Berdasarkan Variabel Pencegahan Terjadinya Stunting," *Prism. Pros. Semin. Nas. Mat.*, vol. 7, pp. 702–711, 2024, [Online]. Available: <https://proceeding.unnes.ac.id/prisma/article/view/3021/2484>
- [12] C. C. Candra, F. Yanuar, and D. Devianto, "Pengelompokan Kabupaten/Kota Di Indonesia Berdasarkan Masalah Gizi Balita Dengan Menggunakan Metode Two Step Cluster Dan Ensemble Clustering," *J. Lebesgue J. Ilm. Pendidik. Mat. Mat. dan Stat.*, vol. 4, no. 2, pp. 1370–1375, 2023, doi: 10.46306/lb.v4i2.413.
- [13] Y. Wang, C. Shan, Y. Zhang, L. Ding, J. Wen, and Y. Tian, "Early Recognition of the Preference for Exclusive Breastfeeding in Current China: A Prediction Model based on Decision Trees," *Sci. Rep.*, vol. 10, no. 1, pp. 1–8, 2020, doi: 10.1038/s41598-020-63073-w.
- [14] S. H. Sørengaard *et al.*, "Exclusive Breastfeeding Duration and Risk of Childhood Cancers," *JAMA Netw. Open*, vol. 7, no. 3, p. E243115, 2024, doi: 10.1001/jamanetworkopen.2024.3115.
- [15] A. Oliver-Roig, J. R. Rico-Juan, M. Richart-Martínez, and J. Cabrero-García, "Predicting exclusive breastfeeding in maternity wards using machine learning techniques," *Comput. Methods Programs Biomed.*, vol. 221, p. 106837, 2022, doi: 10.1016/j.cmpb.2022.106837.