

PENGELOMPOKAN TINGKAT PENDIDIKAN BERDASARKAN JUMLAH SEKOLAH DI PROVINSI JAWA BARAT MENGGUNAKAN ALGORITMA K-MEANS

Zulia Imami Alfianti*¹

^{1,2} Universitas Bina Sarana Informatika Kabupaten Karawang
Jl. Banten No. 1, Karangpawitan, Karawang
E-mail : *zulia.zim@bsi.ac.id

ABSTRAK- Pendidikan di Indonesia terbagi menjadi tiga jenis yaitu pendidikan formal, pendidikan non formal dan pendidikan Informal. Jarak tempuh sekolah dengan tempat tinggal mempengaruhi tingkat pendidikan sekolah dasar dan menengah. Banyak masyarakat yang tidak melanjutkan sekolah dikarenakan jarak tempuh sekolah dengan tempat tinggal memiliki jarak tempuh yang sangat jauh. Peristiwa ini tentunya menurunkan tingkat indeks pendidikan di provinsi Jawa Barat dan meningkatkan angka kemiskinan. Pada penelitian ini akan dilakukan pengklusteran wilayah berdasarkan jumlah sekolah dasar dan menengah untuk menentukan indeks pendidikan di provinsi Jawa Barat menggunakan algoritma *K-Means*. Pengklusteran dilakukan untuk membagi kabupaten/kota di Provinsi Jawa Barat Menjadi tiga *cluster* yaitu tinggi, sedang, dan menengah berdasarkan jumlah sekolah dasar dan menengah pada provinsi Jawa Barat. Hasil dari penelitian ini yaitu terdapat 44% wilayah yang memiliki indeks pendidikan rendah, 37 % memiliki indeks pendidikan sedang, dan 19 % memiliki indeks pendidikan yang tinggi berdasarkan jumlah sekolah dasar menengah yang tersebar di seluruh kabupaten/kota pada provinsi Jawa Barat.

Kata kunci : Pendidikan; Provinsi Jawa Barat; Algoritma K-Means

1. PENDAHULUAN

Pendidikan di Indonesia diatur melalui Undang-Undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional. Pendidikan di Indonesia terbagi ke dalam tiga jalur utama, yaitu formal, nonformal, dan informal. Pendidikan juga dibagi ke dalam empat jenjang, yaitu anak usia dini, dasar, menengah, dan tinggi [1]. Pendidikan merupakan aspek penting dalam kehidupan karena terkait dengan kualitas sumber daya manusia sebagai kunci pertumbuhan Negara. Negara Indonesia menempati peringkat 6 terbawah memiliki kualitas pendidikan yang rendah yaitu ada di peringkat 55 dari 73 Negara ASEAN.

Indeks pendidikan di Provinsi Jawa Barat sangat perlu diperhatikan mengingat beberapa waktu ini banyaknya siswa yang tidak melanjutkan sekolah dikarenakan jarak tempuh sekolah yang sulit dijangkau [2]. Peristiwa ini tentunya menurunkan tingkat sumber daya manusia yang ada di provinsi Jawa Barat dan meningkatkan angka kemiskinan. Untuk mengentas masalah SDM di Indonesia, transformasi pendidikan merupakan suatu keniscayaan karena dengan ini pendidikan manusia Indonesia seutuhnya dapat terlaksana [3]. penurunan jumlah sekolah pada tahun ajaran 2020/2021 secara umum terjadi pada sekolah swasta terutama pada jenjang pendidikan menengah [4]

Beberapa wilayah bahkan masih banyak yang minim pengetahuan dikarenakan kurangnya kesadaran akan pentingnya pendidikan. Dalam penelitian ini penulis akan mengelompokkan wilayah berdasarkan indeks pendidikan di provinsi Jawa Barat menggunakan metode K-Means untuk

mengetahui wilayah yang tingkat pendidikannya tergolong rendah yang tentunya membutuhkan perhatian khusus oleh pemerintah khususnya Dinas Pendidikan Provinsi Jawa Barat.

Data mining yang biasa disebut sebagai Knowledge Discovery in Database (KDD) merupakan kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola hubungan dalam himpunan data yang berukuran besar [5]. Hasil perhitungan dari data mining dapat digunakan dalam pengambilan keputusan. Teknik dalam data mining yang digunakan untuk melakukan pengelompokan sejumlah data atau objek ke dalam cluster (grup) disebut dengan *clustering* yang memungkinkan pengelompokan data jumlah sekolah dasar dan menengah di provinsi Jawa Barat.

2. ISI PENELITIAN

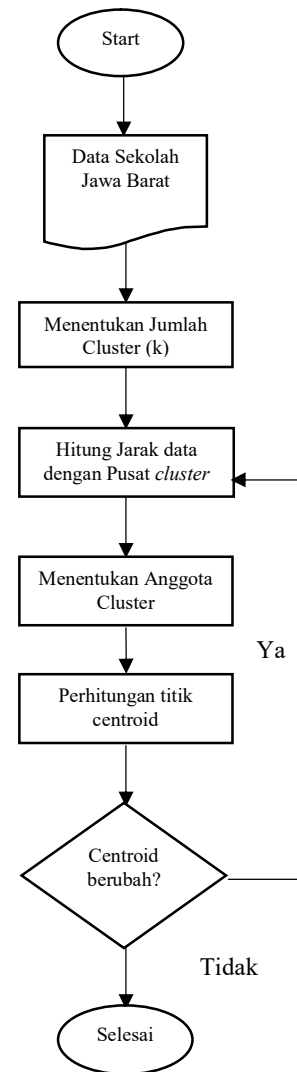
2.1 Metode Penelitian

Penelitian dilakukan dengan menggunakan metode klasifikasi K-Means. K-means merupakan salah satu metode clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster. Metode ini mempartisi data ke dalam cluster sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda di kelompokkan ke dalam cluster yang lain [6]. Data yang digunakan dalam penelitian ini adalah data Referensi Pendidikan (Jumlah sekolah dasar dan menengah dari Kemendikbud di Provinsi Jawa Barat yaitu <https://referensi.data.kemdikbud.go.id> data yang

digunakan adalah data referensi pendidikan yang direkap pada tanggal 7 Desember 2021. Data mining disebut juga Knowledge Discovery in Database (KDD) melakukan ekstraksi informasi dari tumpukan data. Proses pencarian informasi tersebut akan menemukan pola kecenderungan dari data yang kemudian hasil dari penambangan (mining) dapat menjadi informasi yang mudah dipahami [7]. Analisis cluster merupakan salah satu teknik data mining yang bertujuan untuk mengidentifikasi sekelompok objek yang mempunyai kemiripan karakteristik tertentu yang dapat dipisahkan dengan kelompok objek lainnya, sehingga objek yang berada dalam kelompok yang sama relatif lebih homogen daripada objek yang berada pada kelompok yang berbeda [8]. K-Means yaitu salah satu dari metode pengelompokan data nonhierarki (sekatan) yang dapat mempartisi data kedalam bentuk dua kelompok ataupun lebih [9]. K-Means clustering merupakan sebuah konstanta dari sejumlah cluster yang diinginkan, sedangkan Means atau dapat didefinisikan sebagai cluster adalah suatu nilai rata-rata dari sekumpulan populasi data [10]. Metode KMeans digunakan untuk mengelompokkan data ke dalam beberapa kelompok, data dalam satu kelompok mempunyai karakteristik yang sama dan mempunyai karakteristik yang berbeda dengan data dari kelompok yang lain [11]. Metode ini berusaha meminimalkan variasi antar data yang ada di dalam suatu *cluster* dan memaksimalkan variasi data yang ada di *cluster* lainnya.

2.2. Hasil dan Pembahasan

Proses pengelompokan indeks pendidikan daerah berdasarkan jumlah sekolah dasar dan menengah di Provinsi Jawa Barat menggunakan algoritma K-Means. Langkah pertama yang dilakukan adalah menentukan jumlah cluster yang ditentukan yaitu jumlah cluster harus lebih sedikit dari jumlah data yang akan diteliti. Selanjutnya, pilih titik centroid secara acak yang nantinya titik ini akan menjadi centroid dari masing-masing cluster. Kemudian hitung jarak dan alokasikan masing-masing data ke centroid terdekat. Setelah itu, tentukan centroid baru dari data yang ada di masing-masing cluster. Apabila masih ada data yang berpindah cluster atau ada perubahan nilai centroid, maka kembali hitung jarak dan alokasikan kembali data ke masing-masing centroid. Jika tidak ada perpindahan cluster atau perubahan nilai centroid maka hentikan proses clustering. Langkah-langkah dalam metode K-Means dari data yang akan diteliti digambarkan dalam flowchart yang dapat dilihat pada Gambar 1.



Gambar 1. Langkah-Langkah Pengklusteran Menggunakan Metode K-Means

Langkah pertama yang dilakukan pada proses clustering menggunakan metode K-Means adalah menentukan jumlah cluster awal secara bebas yaitu 3 cluster dan centroid

dari data 30 Kecamatan di Kabupaten Karawang yang ditandai dengan K1 sampai

dengan K30. Ditentukan 3 titik centroid awal yaitu K17, K18, K27 diambil dari jumlah kasus tinggi, sedang/menengah, dan rendah. Data sekolah dasar dan menengah pada provinsi jawa barat di dapat dilihat pada Tabel 1.

Tabel 1. Data Sekolah Dasar dan Menengah Kabupaten Provinsi Jawa Barat pada Tanggal 04 Januari 2022

Kota/Kab.	SD	SMP	SMA	SMK
Bogor	2,499	1,099	307	361
Sukabumi	1,563	669	213	163
Cianjur	1,490	556	166	186
Bandung	1,656	576	233	137
Sumedang	670	203	57	89
Garut	1,893	734	260	176
Tasikmalaya	1,315	505	169	135
Ciamis	914	247	84	68
Kuningan	751	177	58	41
Majalengka	759	208	57	55
Cirebon	1,109	342	114	109
Indramayu	1,045	311	86	137
Subang	1,005	252	80	109
Purwakarta	479	176	66	57
Karawang	1,098	269	73	115
Bekasi	1,248	531	168	190
Bandung Barat	884	330	134	100
Pangandaran	345	88	24	31
Kota Bandung	549	310	166	121
Kota Bogor	343	175	73	103
Kota Sukabumi	149	74	38	29
Kota Cirebon	183	61	34	27
Kota Bekasi	852	387	144	141
Kota Depok	580	329	98	125
Kota Cimahi	138	62	24	23
Kota Tasikmalaya	288	136	64	51
Kota Banjar	110	42	13	17

Sumber: <https://referensi.data.kemdikbud.go.id/>

Dari data jumlah sekolah 27 Kabupaten/kota di provinsi jawa barat tersebut ditentukan tiga cluster yaitu data ke-1, 8, dan 27 sebagai titik pusat centroid awal yang dipilih berdasarkan data jumlah sekolah pada masing-masing kota/kabupaten pada provinsi Jawa Barat sehingga didapatkan satu kabupaten yang memiliki jumlah sekolah dasar dan menengah yang terbanyak, satu kabupaten/kota yang mempunyai jumlah sekolah sedang/menengah, dan satu kabupaten/kota yang memiliki jumlah sekolah terendah dari 27 kabupaten/kota pada provinsi jawa barat yang ditetapkan sebagai centroid awal. Centroid awal yang sebelumnya sudah ditentukan dapat dilihat pada tabel 2.

Tabel 2. Hasil Pemilihan Titik *Centroid* Awal

Centroid Awal	Atribut 1	Atribut 2	Atribut 3	Atribut 4
C1	2,499	1,099	307	361
C2	914	247	84	68
C3	110	42	13	17

Setelah ditentukan *centroid* awal, langkah selanjutnya adalah menghitung jarak antara *centroid* awal dengan titik *centroid* menggunakan rumus sebagai berikut:

$$d_{(a,b)} = \sqrt{(2.499 - 2.499)^2 + (1.099 - 1.099)^2 + (307 - 307)^2 + (361 - 361)^2}$$

$$d_{(a,b)} = \sqrt{(914 - 2.499)^2 + (247 - 1.099)^2 + (84 - 307)^2 + (68 - 361)^2}$$

$$d_{(a,b)} = \sqrt{(110 - 2.499)^2 + (42 - 1.099)^2 + (13 - 307)^2 + (17 - 361)^2}$$

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

Tabel 3. Hasil Perhitungan Jarak Cluster dengan Titik Centroid

Kode	C1	C2	C3
K1	0	1836.76	2651.29
K2	1053.10	790.53	1601.76
K3	1167.66	669.25	1490.15
K4	1019.72	828.11	1654.71
K5	2069.91	250.28	588.76
K6	732.73	1112.76	1935.00
K7	1350.85	488.95	1305.62
K8	1836.76	0	834.31
K9	2017.42	181.31	657.04
K10	1994.39	162.61	672.41
K11	1614.28	222.78	1051.98
K12	1683.47	161.31	983.01
K13	1750.56	100.01	926.32
K14	2254.51	441.26	398.15
K15	1663.42	191.49	1020.23
K16	1391.46	462.76	1260.20
K17	1815.93	106.36	838.77
K18	2418.84	594.98	240.12
K19	2121.90	383.04	546.59
K20	2371.37	576.68	288.05
K21	2599.17	786.63	57.56
K22	2574.37	757.05	78.93
K23	1815.08	179.92	837.92
K24	2091.61	348.89	567.58
K25	2616.10	801.26	36.61
K26	2443.57	636.30	210.42
K27	2651.29	834.31	0

Tabel 4. Keanggotaan Cluster Hasil Perhitungan Centroid pada Iterasi 1

Centroid	Keanggotaan Cluster
C1	K1 K6
C2	K2 K3 K4 K5 K7 K8 K9 K10 K11 K12 K13 K15 K16 K17 K19 K23

	K24
C3	K14 K18 K20 K21 K22 K25 K26 K27

Tabel 5. Nilai Centroid dari Hasil Perhitungan Iterasi 1

Centroid Awal	Atribut 1	Atribut 2	Atribut 3	Atribut 4
C1	2196	916.5	283.5	268.5
C2	1028.7	364.82	123.52	118.88
C3	254.37	101.75	42	42.25

Dari hasil perhitungan untuk mencari jarak terpendek pada Tabel 3 didapatkan jarak terpendek antara data dengan centroid sehingga didapatkan keanggotaan cluster yaitu C1 sebanyak 8, C2 sebanyak 8, dan C3 sebanyak 14, dapat dilihat pada Tabel 4. Langkah selanjutnya yaitu dengan kembali menghitung nilai rata-rata anggota cluster sehingga didapatkan nilai centroid iterasi 1 yang dapat dilihat pada Tabel 5: Hasil yang didapatkan di iterasi 1 ternyata ada perubahan nilai centroid, maka diharuskan melakukan perhitungan kembali menggunakan rumus euclidean distance untuk menentukan jarak antara centroid dengan titik centroid yang hasilnya terdapat pada Tabel 6, kemudian dipilih kembali centroid mana dari masing-masing kode yang memiliki jarak terpendek, hasil perhitungan tersebut dapat dilihat pada Tabel 6

Tabel 6. Nilai Centroid Hasil

Kode	C1	C2	C3
K1	366.36	1671.24	2490.91
K2	691.40	622.85	1441.55
K3	805.61	505.61	1330.09
K4	653.74	671.12	1494.96
K5	1709.17	400.21	430.58
K6	366.36	951.41	1774.89
K7	988.14	322.39	1145.54
K8	1473.68	176.61	677.22
K9	1654.53	350.36	502.55
K10	1632.12	325.33	516.06
K11	1251.31	84.59	893.16
K12	1322.01	69.99	824.49
K13	1388.12	123.62	769.35
K14	1894.32	587.34	238.24
K15	1301.05	128.65	863.67
K16	1032.86	287.64	1099.66
K17	1454.67	150.39	678.47
K18	2058.24	749.47	94.08
K19	1765.22	484.69	389.55
K20	2013.73	713.46	133.68
K21	2240.01	934.80	109.84
K22	2214.63	907.73	83.97

K23	1456.85	180.62	677.25
K24	1735.40	450.89	409.46
K25	2256.79	950.87	125.76
K26	2084.49	780.48	53.51
K27	2291.84	985.30	160.91

Tabel 7. Nilai Centroid yang Dihasilkan pada Iterasi 2

Centroid Awal	Atribut 1	Atribut 2	Atribut 3	Atribut 4
C1	2016	803	266.66	224.66
C2	1050.21	356.21	114.5	117
C3	316.4	145.3	60	58.4

Berdasarkan hasil perhitungan nilai centroid pada iterasi 2 terdapat perbedaan antara nilai centroid yang dihasilkan pada iterasi 1. Sehingga diperlukan iterasi berikutnya dengan melakukan perhitungan kembali jarak antara data dengan nilai centroid yang terakhir untuk kemudian ditentukan kembali keanggotaan cluster melalui jarak terkecil antara data dengan cluster serta hitung nilai centroid berikutnya

Berikut ini nilai centroid dari iterasi 2 yang dapat dilihat pada Tabel 7. Hasil yang didapatkan untuk menentukan jarak Antara centroid dengan titik centroid menggunakan rumus euclidean distance pada iterasi 3 yang hasilnya dijabarkan pada Tabel 8, kemudian dipilih centroid mana dari masing-masing kode yang memiliki jarak terpendek, hasil perhitungan tersebut yang dapat dilihat pada Tabel 8.

Tabel 8. Hasil Perhitungan Jarak Cluster dengan Titik Centroid Sebelumnya Untuk Menentukan Centroid pada Iterasi 3

Kode	C1	C2	C3
K1	584.05	1657.49	2413.68
K2	479.42	610.41	1364.77
K3	591.02	490.65	1254.40
K4	435.83	655.53	1419.90
K5	1494.68	414.88	359.59
K6	149.34	936.83	1698.84
K7	773.16	309.10	1069.73
K8	1257.55	183.88	606.74
K9	1438.53	361.40	436.10
K10	1416.62	337.52	447.04
K11	1035.30	61.00	819.98
K12	1106.90	57.30	751.77
K13	1172.15	118.99	698.93
K14	1680.43	603.91	165.58
K15	1085.08	107.77	793.45
K16	821.42	279.03	1022.55
K17	1240.28	170.24	602.90
K18	1843.87	764.75	78.40

K19	1554.35	505.98	310.45
K20	1801.56	731.37	61.21
K21	2026.74	951.53	185.62
K22	2000.91	924.00	162.98
K23	1244.98	204.16	599.30
K24	1524.84	471.35	330.31
K25	2043.40	967.32	203.25
K26	1871.39	797.72	31.04
K27	2078.32	1001.51	239.15

Setelah dilakukannya perhitungan jarak data dengan titik centroid, didapatkan hasil perhitungan centroid pada iterasi 3. Berikut ini nilai centroid dari iterasi 3 yang dapat dilihat pada Tabel 9.

Tabel 9. Nilai Centroid yang Dihasilkan pada Iterasi 3

Centroid Awal	Atribut 1	Atribut 2	Atribut 3	Atribut 4
C1	1902.75	769.5	253.25	209.25
C2	1039.16	342.91	111.08	115.5
C3	348.54	150.54	59.72	61.18

Tabel 10. Hasil Perhitungan Jarak Cluster dengan Titik Centroid Sebelumnya Untuk Menentukan Centroid pada Iterasi 4

Kode	C1	C2	C3
K1	700.00	1673.74	2382.23
K2	359.56	627.19	1333.24
K3	473.38	506.59	1222.36
K4	322.42	670.92	1387.88
K5	1376.07	399.35	326.90
K6	50.06	952.79	1667.04
K7	654.23	325.71	1027.81
K8	1139.83	166.90	574.17
K9	1320.61	344.87	403.83
K10	1298.36	321.37	414.51
K11	917.73	70.20	787.51
K12	989.52	46.30	719.18
K13	1055.37	102.18	666.27
K14	1561.26	589.15	133.12
K15	969.27	101.86	760.77
K16	702.29	296.27	991.00
K17	1121.23	158.14	570.91
K18	1724.92	749.38	78.17
K19	1434.98	494.36	283.68
K20	1682.26	717.25	50.53
K21	1907.40	936.76	217.22
K22	1881.70	908.99	193.01
K23	1125.60	196.74	568.19
K24	1405.31	459.66	301.58
K25	1924.09	952.44	234.31
K26	1752.02	783.22	63.23
K27	1959.05	986.52	269.85

Tabel 11. Nilai Centroid yang Dihasilkan pada Iterasi 4

Centroid Awal	Atribut 1	Atribut 2	Atribut 3	Atribut 4
C1	1820.2	726.8	235.8	204.6
C2	998.18	323.54	106.09	109.09
C3	348.54	150.54	59.72	61.18

Tabel 12. Hasil Perhitungan Jarak Cluster dengan Titik Centroid Sebelumnya Untuk Menentukan Centroid pada Iterasi 5

Kode	C1	C2	C3
K1	792.98	1719.76	2382.23
K2	267.84	672.82	1333.24
K3	378.71	552.65	1222.36
K4	232.98	716.47	1387.88
K5	1281.66	353.62	326.90
K6	82.19	998.66	1667.04
K7	560.11	371.38	1037.81
K8	1045.51	122.97	574.17
K9	1226.31	299.20	403.83
K10	1204.01	275.48	414.51
K11	823.31	112.62	787.51
K12	894.89	59.42	719.18
K13	960.93	76.45	666.27
K14	1467.24	534.72	133.12
K15	875.03	118.61	760.77
K16	608.73	340.33	991.00
K17	1027.24	118.07	570.91
K18	1630.72	703.53	78.17
K19	1342.21	453.51	283.68
K20	1588.53	672.65	50.53
K21	1813.56	891.31	217.22
K22	1787.72	863.35	193.01
K23	1032.15	166.88	568.19
K24	1312.12	418.59	301.58
K25	1830.18	906.90	243.31
K26	1658.24	738.02	63.23
K27	1865.09	940.89	269.85

Tabel 12. Nilai Centroid yang Dihasilkan pada Iterasi 5

Centroid Awal	Atribut 1	Atribut 2	Atribut 3	Atribut 4
C1	1820.2	726.8	235.8	204.6
C2	998.18	323.54	106.09	109.09
C3	348.54	150.54	59.72	61.18

Setelah dilakukan perhitungan, proses iterasi berhenti di iterasi 5 karena nilai centroid iterasi 5 sama dengan nilai centroid iterasi 4. Perbandingan nilai centroid yang berasal dari iterasi awal sampai dengan iterasi 3 tersebut dapat dilihat pada Tabel 13.

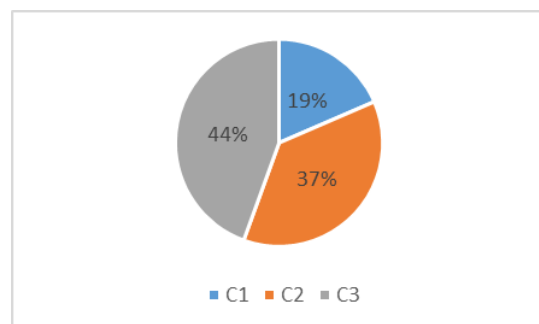
Tabel 13. Perbandingan Nilai Centroid pada Iterasi Awal sampai Iterasi 5

Centroid Awal	Atribut 1	Atribut 2	Atribut 3	Atribut 4
		Iterasi 0		
C1	2,499	1,099	307	361
C2	914	247	84	68
C3	110	42	13	17
		Iterasi 1		
C1	2196	916.5	283.5	268.5
C2	1028.70	364.82	123.52	118.88
C3	245.37	101.75	42	42.25
		Iterasi 2		
C1	2016	803	266.66	224.66
C2	1050.21	356.21	114.5	117
C3	316.4	145.3	60	58.4
		Iterasi 3		
C1	1902.75	769.5	253.25	209.25
C2	1039.15	342.91	111.08	115.5
C3	348.54	150.54	59.72	61.18
		Iterasi 4		
C1	1820.2	726.8	235.8	204.6
C2	998.18	323.54	106.09	59.72
C3	348.54	150.54	59.72	61.18
		Iterasi 5		
C1	1820.2	726.8	235.8	204.6
C2	998.18	323.54	106.09	59.72
C3	348.54	150.54	59.72	61.18

Setelah dilakukan perbandingan nilai centroid pada Tabel 13, proses iterasi berhenti pada iterasi 5 karena nilai centroid pada iterasi 5 sama dengan nilai centroid pada iterasi 4. Sehingga didapatkan keanggotaan cluster seperti ditampilkan pada Tabel 14.

Tabel 14. Keanggotaan Cluster Hasil Perhitungan Centroid pada Iterasi 5

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



Gambar 2. Diagram Hasil Pengklasteran Index Pendidikan Sekolah dasar dan menengah Wilayah Provinsi Jawa Barat

Pada diagram lingkaran diatas yang dihasilkan dari proses clustering terdapat 19% yaitu 5 kabupaten/kota yang memiliki tingkat pendidikan tinggi, 37% terdiri dari 10 kabupaten/kota memiliki tingkat pendidikan sedang/menengah karena berada diantara persentasi tingkat tinggi dan rendah, dan 44% terdiri dari 12 kelurahan memiliki tingkat pendidikan rendah.

3. KESIMPULAN

Dari hasil clustering jumlah sekolah dasar dan menengah di provinsi jawa barat diatas, maka untuk lebih jelasnya penulis gambarkan pada sebuah diagram lingkaran yang ditampilkan pada Gambar 2. Hasil perhitungan menggunakan metode K-Means pada diagram lingkaran yang terdapat pada Gambar 2, dari 27 kabupaten/kota di wilayah Provinsi Jawa Barat yang telah melewati proses clustering menggunakan metode K-Means dibagi menjadi tiga cluster indeks pendidikan yang dikategorikan menjadi tiga wilayah tingkat pendidikan yang memiliki tingkat pendidikan tinggi memiliki jumlah sekolah terbanyak, wilayah yang memiliki tingkat pendidikan sedang/menengah yang memiliki jumlah sekolah dasar dan menengah sedang, dan wilayah yang memiliki tingkat pendidikan rendah memiliki jumlah sekolah dasar dan menengah berjumlah sedikit.

4. PENUTUP

Setelah dilakukan proses pengelompokan wilayah untuk menentukan tingkat pendidikan di Provinsi Jawa Barat dengan menggunakan metode K-Means dapat disimpulkan bahwa dari 27 kabupaten/kota yang ada di Provinsi Jawa Barat 44% wilayah termasuk kedalam kategori tingkat pendidikan yang rendah, 37% wilayah termasuk kedalam kategori tingkat pendidikan sedang, dan 19% wilayah termasuk kedalam kategori tingkat pendidikan tinggi berdasarkan jumlah sekolah dasar dan menengah di provinsi Jawa Barat.

Untuk mengembangkan penelitian ini, dibutuhkan penelitian lanjutan metode clustering terhadap tingkat pendidikan yang lebih luas yaitu pada 34 provinsi di Indonesia yang dilakukan secara lebih valid dengan ditetapkan nilai centroid terbaik untuk mengetahui pengelompokan wilayah berdasarkan tingkat pendidikan di Indonesia.

DAFTAR PUSTAKA

- [1] N. M. Hasanah, "Penyelenggaraan Jalur Pendidikan Formal dan Nonformal (Studi Kasus di PAUD Terpadu Salsabila Al-Muthi'in Yogyakarta)," *Desember*, vol. 1, no. 2, pp. 84-97, 2019.
- [2] R. Nur Hayati, Supriadi, "ANALISIS FAKTOR PENYEBAB ANAK TIDAK MELANJUTKAN PENDIDIKAN KE

- SEKOLAH MENENGAH ATAS Nur Hayati, Supriadi, Rustiyarso Program Studi Pendidikan Sosiologi FKIP UNTAN Pontianak,” no. 1, pp. 1–11, 2011.
- [3] H. Widodo, “Potret Pendidikan Di Indonesia Dan Kesiapannya Dalam Menghadapi Masyarakat Ekonomi Asia (Mea),” *Cendekia J. Educ. Soc.*, vol. 13, no. 2, p. 293, 2016, doi: 10.21154/cendekia.v13i2.250.
- [4] “Statistik Pendidikan 2021 Badan Pusat Statistik,” *Badan Pus. Stat.*, 2021.
- [5] 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.
- [6] H. Sy, Rismayani, and A. Syam, “Data Mining Menggunakan Algoritma K-Means Pengelompokan Penyebaran Diare Di Kota Makassar,” *SISITI Semin. Ilm. Sist. Inf. dan Teknol. Inf.*, vol. 8, no. 1, pp. 73–82, 2019.
- [7] 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.
- [8] T. Suprawoto, “Klasifikasi data mahasiswa menggunakan metode k- means untuk menunjang pemilihan strategi pemasaran,” vol. 1, no. 1, pp. 12–18, 2016.
- [9] M. H. Adiya and Y. Desnelita, “Jurnal Nasional Teknologi dan Sistem Informasi Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru,” vol. 01, pp. 17–24, 2019.
- [10] dkk Sugiono, “Pengelompokan Perilaku Mahasiswa Pada Perkuliahan E-Learning dengan K-Means Clustering,” *J. Kaji. Ilm.*, vol. 19, no. 2, pp. 126–133, 2019.
- [11] Z. I. Alfianti, M. A. Azis, A. Fauzi, F. T. Informasi, and U. B. Saranainformatika, “JurnalMantik,” vol. 4, no. 4, pp. 2336–2341, 2021.

Contact person Author: Zulia Imami Alfianti
Hp :085894753748