

PENERAPAN METODE *TWO-STEP CLUSTER* UNTUK PENGELOMPOKAN DESA BERDASARKAN KEPADATAN PENDUDUK

Riri Narasati¹, Sandy Eka Permana² Kaslani³ Edi Tohidi⁴ Iin⁵

Program Studi Teknik Informatika, STMIK IKMI Cirebon, Kota Cirebon Indonesia¹
Program Studi Manajemen Informatika, STMIK IKMI Cirebon, Kota Cirebon Indonesia²
Program Studi Manajemen Informatika, STMIK IKMI Cirebon, Kota Cirebon Indonesia³
Program Studi Rekayasa Perangkat Lunak, STMIK IKMI Cirebon, Kota Cirebon Indonesia⁵

Email : narasati5600@gmail.com¹, sandyekapermana@ikmi.ac.id², kaslani@gmail.com³, edy.tohidi1234@gmail.com⁴,
isholihin87@gmail.com⁵

Email Penulis Korespondensi : narasati5600@gmail.com¹

Abstract-Population density has a significant impact on the quality of life of the city. Due to the high population density, many problems arise. Due to the high population density, many problems arise. Population is continuously affected by the number of babies born (population growth), but at the same time it is reduced by the number of deaths that occur in all age groups. Population growth is caused by four factors, namely births, deaths, in-migration and out-migration. When immigrants enter in large numbers, the number of residents in an area or city will increase and when immigrants leave increases then the opposite is the number of residents in an area or city will decrease. This study aims to analyze the level of population density in villages in Majalengka Regency. Majalengka Regency is one of the districts in West Java. As the population of Majalengka increases every year, the population density in Majalengka Regency is quite dense. In this study, villages in Majalengka Regency were clustered using the two-step cluster method. Two-Step Cluster is a way to handle continuous and categorical variables by using Bayesian Information Criterion (BIC) indicators, the result of the optimal number of clusters obtained 4 clusters. In this cluster there is a cluster which is the best cluster, which is in cluster 3 with the number of villages, namely 15 villages with a population of 6,862 people. The clustering test was carried out using the silhouette method and showed that the resulting cluster had a silhouette value which got a sufficient value of 0.4. This indicates that the clusters formed have been well clustered.

Keywords : cluster, density, Two Step Cluster

Abstrak- Kepadatan penduduk memiliki dampak yang signifikan terhadap kualitas hidup kota. Karena kepadatan penduduk yang tinggi banyak masalah yang muncul. Penduduk terus menerus dipengaruhi oleh jumlah bayi yang lahir (pertumbuhan penduduk), tetapi pada saat yang sama dikurangi dengan jumlah kematian yang terjadi pada semua kelompok umur. Pertumbuhan penduduk disebabkan oleh empat faktor yaitu kelahiran, kematian, imigrasi masuk, dan migrasi keluar. Ketika imigran masuk berjumlah banyak maka jumlah penduduk di suatu wilayah atau kota akan bertambah dan ketika imigran keluar bertambah maka sebaliknya yaitu jumlah penduduk pada suatu wilayah atau kota akan berkurang. Penelitian ini bertujuan untuk menganalisis tingkat kepadatan penduduk pada desa di Kabupaten Majalengka. Kabupaten Majalengka merupakan salah satu kabupaten yang ada di Jawa Barat. Seiring bertambahnya jumlah penduduk Majalengka setiap tahunnya, kepadatan penduduk di Kabupaten Majalengka cukup padat. Dalam penelitian ini, desa-desa di Kabupaten Majalengka di *cluster* menggunakan

metode *Two-Step Cluster*. *Two-Step Cluster* adalah cara untuk menangani variabel kontinu dan kategorik dengan menggunakan indikator *Bayesian Information Criterion (BIC)*, hasil dari jumlah *cluster* optimal diperoleh 4 *cluster*. Pada *cluster* tersebut terdapat *cluster* yang merupakan *cluster* terbaik yaitu terdapat pada *cluster* 3 dengan jumlah desa yaitu 15 desa dengan jumlah penduduk 6.862 Jiwa. Pengujian *clustering* dilakukan dengan metode *silhouette* dan menunjukkan bahwa *cluster* yang dihasilkan memiliki nilai *silhouette* yang mendapatkan nilai cukup yaitu 0,4. Hal ini menunjukkan bahwa *cluster* yang terbentuk telah ter-*cluster* dengan baik.

Kata Kunci : cluster, kepadatan, *Two-Step Cluster*

I. PENDAHULUAN

Kepadatan penduduk merupakan suatu keadaan yang disebut lebih padat ketika jumlah penduduk dalam suatu batas ruang tertentu meningkat dibandingkan dengan luas ruang tersebut. Penduduk dapat diartikan sebagai organisme tunggal dari jenis individu yang sama yang hidup di suatu daerah dengan batas tertentu. Pertumbuhan penduduk Indonesia masih menjadi masalah utama terutama di perkotaan. Tentunya seiring dengan pertumbuhan jumlah penduduk, kita perlu menambah berbagai fasilitas sosial yang menunjang kehidupan masyarakat. (Kurniawati & Sugiyanto, 2021). Pertumbuhan penduduk merupakan keseimbangan dinamis antara pertumbuhan dan penurunan penduduk. Hal ini tidak hanya tergantung pada jumlah bayi yang lahir, berdasarkan jenis kelamin (gender), tetapi juga pada jumlah kematian pada semua kelompok umur. (Silastri et al., 2017)

Penduduk yang telah terdaftar sebagai penduduk dengan maksud tinggal dari tempat asalnya ke tempat tujuan selama enam bulan dapat ditetapkan sebagai penduduk tetap. Penduduk yang mengunjungi tempat tujuan secara teratur dan tidak bermaksud untuk tinggal di luar batas tempat asalnya disebut penduduk tidak tetap. (Budianti, 2021). Migrasi juga merupakan salah satu faktor yang mempengaruhi pertumbuhan penduduk. Serbuan penduduk ke daerah tujuan dan perpindahan penduduk dari daerah asalnya juga menyebabkan bertambah atau berkurangnya jumlah penduduk

di daerah tersebut. pertumbuhan penduduk yang tinggi merupakan salah satu prioritas yang harus segera ditangani. Namun, pertumbuhan penduduk yang tidak memadai juga berdampak negatif terhadap pembangunan daerah. Pertumbuhan penduduk sangat mungkin rendah di daerah dengan kelahiran sangat rendah, tetapi jumlah kematian sangat tinggi dan jumlah orang yang meninggalkan daerah karena kondisi tertentu sangat tinggi. (Ainy et al., 2019). Jika jumlah kelahiran melebihi jumlah kematian maka pertumbuhan penduduk akan meningkat. Jika jumlah kedatangan lebih besar dari jumlah keberangkatan, hal ini akan menyebabkan pertumbuhan penduduk di daerah tersebut. (Rumengan et al., 2019).

Penelitian yang dilakukan oleh (Rumengan et al., 2019) menjelaskan bahwa pertumbuhan penduduk yang begitu besar masih menjadi masalah, terutama di satu daerah. Menurut (Silastri et al., 2017) Pertumbuhan penduduk merupakan keseimbangan yang dinamis antara kekuatan-kekuatan yang menambah dan kekuatan-kekuatan yang mengurangi jumlah penduduk. Secara terus menerus penduduk akan dipengaruhi oleh jumlah bayi yang lahir (menambah jumlah penduduk), tetapi secara bersamaan pula akan dikurangi oleh jumlah kematian. Sementara itu migrasi juga berpengaruh terhadap kepadatan penduduk. Migrasi yang tergolong menjadi 2 yaitu Imigran (pendatang) akan mempengaruhi jumlah penduduk dan Emigran akan mengurangi jumlah penduduk.

Kepadatan penduduk berdampak signifikan terhadap kualitas hidup masyarakat, karena menimbulkan banyak masalah yang berkaitan dengan masalah kepadatan penduduk seperti kemiskinan, perumahan dan lapangan pekerjaan. Pembahasan tentang kualitas hidup tidak mudah, dan karena ada indikator objektif dan subjektif kualitas hidup (pendidikan, kesehatan, pekerjaan) dan indikator subjektif (relatif), pemahaman kualitas hidup yang diterima secara universal tidak ada definisi. Daerah padat penduduk sering menimbulkan masalah sosial seperti kemiskinan, pendidikan yang buruk, dan masalah kecemburuan sosial lainnya (Triyastuti, 2019). Dari permasalahan yang telah diuraikan, penelitian ini akan mengidentifikasi kasus kepadatan penduduk pada suatu wilayah yaitu Desa di Kabupaten Majalengka dengan pengelompokan jumlah kepadatan penduduk dengan menggunakan metode *Two-Step Cluster*. Langkah awal yang dilakukan untuk penelitian ini adalah mengumpulkan dataset. Dataset yang digunakan pada penelitian ini yaitu dataset publik yang didapatkan dari website kemendagri sebagai data yang dipakai dalam proses pengelompokan desa berdasarkan kepadatan penduduk dengan menggunakan metode *Two-Step Cluster*. Hasil evaluasi dari pengelompokan desa berdasarkan kepadatan penduduk dengan metode *Two-Step Cluster* ini untuk mengetahui hasil *cluster* dengan jumlah kepadatan penduduk di Kabupaten Majalengka.

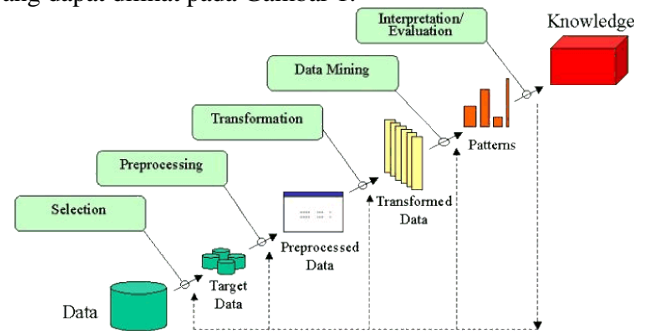
Berdasarkan permasalahan yang dijelaskan dan didukung oleh penelitian sebelumnya, terdapat beberapa teknik pengelompokan terhadap kepadatan penduduk diantaranya *Two-Step Cluster*. *Two-Step Cluster* adalah metode yang dapat menangani variabel kategorik dan kontinu serta memiliki objek dengan jumlah yang relatif besar. Terdapat dua proses

yang dapat dilakukan yaitu *pre-clustering* dan *cluster optimal*. Proses *pre-clustering* memeriksa data dari 0 sampai N satu per satu untuk menentukan dimana data tersebut akan termasuk kedalam *cluster* yang sudah ada atau akan membentuk *cluster* baru sesuai dengan kriteria jarak. Beberapa metode pengukuran yang dapat digunakan untuk menentukan jarak, diantaranya jarak *euclidean* dan *log-likelihood*. (Setiyawati & Bangkalang, 2020). Sedangkan untuk menentukan jumlah *cluster* optimal dapat ditentukan secara otomatis. (Setyawan & Pratiwi, 2019). Data hasil *pre-clustering* digunakan sebagai input untuk pencarian *cluster*. *Two-Step Cluster* menggunakan metode *hierarkis* dalam proses *clustering*. (Setiyawati & Bangkalang, 2020). Untuk menentukan jumlah *cluster* yang akan dibentuk, penelitian ini menggunakan indikator *Schwartz Bayesian Criterion* (BIC). BIC digunakan untuk mencari estimasi jumlah *cluster* dengan perbandingan jarak untuk mendapatkan hasil *cluster* dengan jumlah kepadatan penduduk di Kabupaten Majalengka.

Tujuan dari penelitian ini untuk mengetahui proses pengelompokan Desa berdasarkan kepadatan penduduk di Kabupaten Majalengka dengan penerapan algoritma *Two-Step Cluster* dan menggunakan proses *Schwartz's Bayesian Criterion* (BIC). Pengelompokan tersebut dikelompokkan berdasarkan jumlah laki-laki, jumlah perempuan, jumlah penduduk, jumlah kepemilikan kartu keluarga (KK) dan kepadatan penduduk pada Tahun 2020-2021. Dan untuk menganalisis dan memahami pola kepadatan penduduk di berbagai desa. Dengan menerapkan metode *Two-Step Cluster*, penelitian ini dapat membantu mengidentifikasi desa-desa dengan tingkat kepadatan penduduk yang serupa

II. METODOLOGI PENELITIAN

Knowledge Discovery in Databases (KDD) banyak digunakan untuk menemukan informasi tersembunyi dalam database besar. *Knowledge Discovery in Databases* (KDD) membantu untuk mengidentifikasi dan mengamati apa yang ingin ketahui, memproses data menjadi informasi, dan menghubungkannya. pada penelitian ini menggunakan proses *Knowledge Discovery Databases* yang mempunyai 9 langkah yang dapat dilihat pada Gambar 1.



Gambar 1 Alur Penelitian

Berikut adalah penjelasan dari setiap tahapan *Knowledge Discovery Databases* (KDD) :

Data Selection Pemilihan data (seleksi) dari sekumpulan data operasional harus dilakukan sebelum memulai untuk

ketahap ekstraksi *Knowledge Discovery in Databases* (KDD) dimulai. Data terpilih yang digunakan dalam proses data mining disimpan dalam file terpisah dari database produksi.

Preprocessing/cleaning Sebelum menjalankan proses data mining, perlu menjalankan proses preprocessing pada data yang menjadi fokus *Knowledge Discovery in Databases* (KDD). Proses pembersihan data meliputi menghapus data duplikat, memeriksa data yang tidak konsisten, dan memperbaiki kesalahan data, seperti kesalahan cetak (tipografi).

Transformation Dalam tahap ini dilakukan pengkodean sebagai bentuk transformasi terhadap data hasil dari proses cleaning, kemudian data tersebut dilakukan proses data mining. Proses coding dalam penemuan *knowledge* merupakan proses yang tidak baku dan sangat tergantung pada pola dengan jenis informasi yang akan dicari dalam keseluruhan basis data.

Data Mining Pada tahap proses data mining adalah proses pencarian pola atau informasi yang menarik pada data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma penambangan data sangat berbeda. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *Knowledge Discovery in Databases* (KDD) secara keseluruhan.

Interpretation/Evaluation Tahap ini merupakan bagian dari proses *Knowledge Discovery in Databases* (KDD) untuk memeriksa apakah pola atau informasi yang ditemukan tidak sesuai dengan fakta atau hipotesis yang ada. Pola informasi yang dihasilkan dari proses data mining harus disajikan dalam format yang mudah dipahami oleh para pemangku kepentingan.

Data yang digunakan dalam penelitian ini adalah data sekunder dan diperoleh dari *website* kemendagri tentang jumlah kepadatan penduduk. Data yang digunakan sebanyak 332 Desa yang ada di Kabupaten Majalengka dan mencakup 12 atribut berdasarkan jumlah kepadatan Tahun 2020-2021. Dataset yang digunakan adalah data sekunder yaitu data yang diperoleh bukan hasil dari pengumpulan data secara langsung oleh peneliti. Dataset diperoleh dari penyedia data kependudukan dengan halaman *website* (<http://prodeskel.binapemdes.kemendagri.go.id/>).

Teknik Pengumpulan Data Pengumpulan data yang berasal dari hasil observasi yaitu sebagai berikut : Dataset yang diperoleh dari *website* kemendagri tentang jumlah kepadatan penduduk yang ada di wilayah Jawa barat Tahun 2020-2021. Data yang diperoleh sudah berbentuk excel. Data yang diperoleh dengan mengkaji studi literatur berupa aturan-aturan tertulis atau dokumen yang ada kaitannya dengan judul penelitian.

Populas dataset pada penelitian ini sebanyak 332 Desa yang ada di Kabupaten Majalengka dengan 12 atribut diantaranya Provinsi, Kabupaten/Kota, Kecamatan, Status (Desa/Kelurahan), Kode UPM, Desa/Kelurahan, Tahun, Laki-Laki, Perempuan, Jumlah penduduk, Jumlah Kartu keluarga (KK), Kepadatan Penduduk (Jiwa/Km²).

III. HASIL DAN PEMBAHASAN

A. Data

Penelitian pada tahap ini, data yang diperoleh dari *website* kemendagri tentang data kepadatan penduduk. Terdapat 343 Desa di Kabupaten Majalengka dengan 12 atribut diantaranya Provinsi, Kabupaten/Kota, Kecamatan, Status, Kode UPM, Desa/Kelurahan, Tahun, Jumlah laki-laki, jumlah perempuan, Jumlah penduduk, Jumlah Kartu keluarga (KK), Kepadatan (Jiwa/Km²).

ID_NO	KECAMATAN	DESA/KELURAHAN	KODE UPM	DESA	TAHUN	JUMLAH LAKI LAKI	JUMLAH PEREMPUAN	JUMLAH PENDUDUK	JUMLAH KEPEREMPUAN KK	KEPADATAN PENDUDUK
1	LEMAHSLUGUH	DESA	3210012001	CPASUNGGI	2020	115	1149	2299	796	391
2	LEMAHSLUGUH	DESA	3210012002	BOROGOKIL	2021	1917	1912	3869	151	412
3	LEMAHSLUGUH	DESA	3210010003	BANGBAYANG	2021	1129	1048	2177	704	254
4	LEMAHSLUGUH	DESA	3210012004	CIBULAN	2021	832	833	1665	493	277
5	LEMAHSLUGUH	DESA	3210012005	LEMAHPUTIH	2021	2383	2382	4765	1488	514
6	LEMAHSLUGUH	DESA	3210012006	SADAWANGI	2021	2203	2142	4395	1384	417
7	LEMAHSLUGUH	DESA	3210012011	MARGALAYA	2021	2562	2431	4995	1416	437
8	LEMAHSLUGUH	DESA	3210012008	KALAJADINA	2021	1994	1928	3922	1122	349
9	LEMAHSLUGUH	DESA	3210012009	CUHALUH	2021	219	2376	4594	1604	509
10	LEMAHSLUGUH	DESA	3210012013	MEKARALAYA	2021	179	151	33	1026	1401
11	LEMAHSLUGUH	DESA	3210012014	SINARIGALUH	2021	3062	2897	5989	1426	453
12	LEMAHSLUGUH	DESA	3210010010	PANGREKA	2021	2011	2701	4786	179	1261
136	SUMBERALAYA	DESA	320012001	PANGCAPUTAT	2021	1252	1235	2487	853	2
137	SUMBERALAYA	DESA	320012002	BONGAS WETAN	2021	2172	1931	4053	2126	2077
138	SUMBERALAYA	DESA	320012003	BONGAS KULON	2021	1974	2001	3975	1216	165
139	SUMBERALAYA	DESA	320012004	PANALUN LOR	2021	2625	2689	5308	215	1886
140	SUMBERALAYA	DESA	320012005	CIDONGK	2021	1724	1742	3466	887	749
141	SUMBERALAYA	DESA	320012011	LOHOBONG	2021	2492	2569	5061	1877	749
142	SUMBERALAYA	DESA	320012014	PANGKASUHI	2021	1012	985	1997	796	1489
143	SUMBERALAYA	DESA	320012015	GELOM MACTA	2021	1067	1019	2088	646	299

Gambar 2. Data Set

B. Selection

Tahap seleksi ini melengkapi data yang digunakan dalam proses penelitian. Data yang dipilih adalah kepadatan penduduk dan fokus penelitian adalah data desa berdasarkan jumlah penduduk Kabupaten Majalengka. Ada 332 desa/kelurahan di Kabupaten Majalengka. Basis data menggunakan beberapa variabel. Variabel dan tipe data yang digunakan ditunjukkan di bawah ini.

Tabel 1 Variabel data

Variabel	Type Data	Keterangan
Provinsi	<i>Numeric</i>	Provinsi
Kabupaten	<i>String/kontinu</i>	Kabupaten Majalengka
Kecamatan	<i>String/kontinu</i>	Kecamatan di Kabupaten Majalengka
Desa/Kelurahan	<i>String/kontinu</i>	Desa atau kelurahan yang ada di Kabupaten Majalengka
Kode UPM	<i>Numeric</i>	Kode UPM setiap desa
Desa	<i>String/kontinu</i>	Desa yang ada di Kabupaten Majalengka
Tahun	<i>Numeric</i>	Penduduk pada tahun 2020-2021
Jumlah Laki-laki	<i>Numeric</i>	Jumlah laki-laki di Kabupaten Majalengka pada tahun 2020-2021
Jumlah Perempuan	<i>Numeric</i>	Jumlah perempuan di Kabupaten Majalengka pada tahun 2020-2021
Jumlah Penduduk	<i>Numeric</i>	Jumlah penduduk laki-laki dan perempuan di Kabupaten Majalengka
Jumlah KK	<i>Numeric</i>	Jumlah kepemilikan Kartu Keluarga pada tahun 2020-2021
Kepadatan Penduduk	<i>Numeric</i>	Jumlah kepadatan penduduk pada tahun

Berdasarkan variabel-variabel yang terdapat pada tabel di atas pengelompokan desa di Kabupaten Majalengka dengan *Two-Step Cluster* tersebut hanya menggunakan 5 variabel, diantaranya yaitu :

- a) X7 = Jumlah Laki-laki
- b) X8 = Jumlah Perempuan
- c) X9 = Jumlah penduduk
- d) X10 = Jumlah kepemilikan kartu keluarga (KK)
- e) X11 = Kepadatan penduduk

Hasil dari pengelompokan ini akan membentuk menjadi beberapa *cluster* berdasarkan jumlah kepadatan penduduk disetiap desa di Kabupaten Majalengka.

C. Preprocessing

Pada tahap ini, pemeriksaan data dilakukan untuk menentukan apakah data memiliki nilai yang hilang, nama kolom acak, atau nomor yang tidak cocok. Nilai-nilai yang hilang dari informasi tentang jumlah dan tahun kepadatan penduduk. Jika Anda tidak memperbaiki data nilai yang hilang, pemrosesan berikut akan terhambat. Ada beberapa teknik untuk mengatasi missing value yaitu diantaranya dapat mengabaikan atau menghapus data, mengubah data secara manual, atau menggunakan nilai yang dapat mengubah data nilai yang hilang. Setelah verifikasi data penduduk di Kabupaten Majalengka terdapat 343 penduduk namun pada data tersebut terdapat desa yang double dan terdapat tahun yang tidak sesuai sehingga data tersebut dihapus dan dataset Kab.Majalengka yang digunakan menjadi 332 desa yang terdapat di Kabupaten Majalengka seperti yang tertera dibawah ini.

ID	NO	DESA	DESA	TAHUN	JUMLAH LAKI-LAKI	JUMLAH PEREMPUAN	JUMLAH PENDUDUK	JUMLAH KEPEMILIKAN KK	KEPADATAN PENDUDUK
1	LEMAHSUGH	DESA	3210012001	2021	115	119	239	790	191
2	LEMAHSUGH	DESA	3210012003	2021	1937	1932	3869	103	415
3	LEMAHSUGH	DESA	3210012005	2021	1129	1048	2177	704	294
4	LEMAHSUGH	DESA	3210012006	2021	832	853	1685	493	277
5	LEMAHSUGH	DESA	3210012005	2021	2283	2182	4465	1488	514
6	LEMAHSUGH	DESA	3210012006	2021	2233	2142	4395	1384	1332
7	LEMAHSUGH	DESA	3210012011	2021	2962	2433	4995	1416	1337
8	LEMAHSUGH	DESA	3210012008	2021	1994	1928	3922	1122	1508
9	LEMAHSUGH	DESA	3210012010	2021	219	2374	4564	1604	199
10	LEMAHSUGH	DESA	3210012015	2021	179	151	331	1026	1401
11	LEMAHSUGH	DESA	3210012014	2021	3062	2807	5869	1426	453
12	LEMAHSUGH	DESA	3210012012	2021	2811	2755	5566	139	1281
129	SUMBERJAYA	DESA	3210172009	2021	107	2062	2162	760	12541
130	SUMBERJAYA	DESA	3210172009	2021	4302	4008	8311	3469	278
131	SUMBERJAYA	DESA	38602905	2021	1252	1235	2487	853	466
132	SUMBERJAYA	DESA	3210172001	2021	3372	3281	6653	2125	2007
133	SUMBERJAYA	DESA	3210172002	2021	1974	2051	3925	1216	160
134	SUMBERJAYA	DESA	3210172012	2021	2825	2683	5508	215	1586
135	SUMBERJAYA	DESA	3210172010	2021	1724	1742	3466	1087	748
136	SUMBERJAYA	DESA	3210172011	2021	2492	2569	5061	1877	749
137	SUMBERJAYA	DESA	3210172014	2021	1012	985	1997	796	1488
138	SUMBERJAYA	DESA	3210172015	2021	1067	1019	2086	640	295

Gambar 3. Preprocessing

D. Datamining

Metode *clustering* untuk pengelompokan desa berdasarkan data kepadatan penduduk adalah dengan menggunakan metode *Two-Step Clustering*. Metode ini digunakan karena datasetnya besar dan variabel data campurannya adalah angka, orde, dan nama. Tahapan dalam *cluster* dua tahap membuat tahap *pre-cluster* untuk membentuk anak-anak dari kelompok. Fase ini bertujuan untuk meminimalkan jumlah data yang relatif besar. Fase ini menggunakan konsep pengelompokan sekuensial. Dengan pendekatan ini, data diperiksa secara acak satu per satu, dan data secara otomatis diubah menjadi data grup sebelumnya atau membentuk grup baru berdasarkan kriteria jarak.

Hasil cluster yang didalamnya terbagi menjadi 4 cluster dengan jumlah desa yang ada di Kabupaten Majalengka.

Cluster	N	% of Combine	% of total
Cluster 1	61	18,4%	18,4%
Cluster 2	135	40,7%	40,7%
Cluster 3	15	4,5%	4,5%
Cluster 4	121	36,4%	36,4%
Total	332		100.0%

Pada tabel di atas terdapat 332 desa di Kabupaten Majalengka yang terbagi dalam 4 cluster. Setiap cluster memiliki jumlah desa yang berbeda-beda, antara lain cluster 1 terdiri dari 61 desa, cluster 2 terdiri dari 135 desa, cluster 3 terdiri dari 15 desa, dan cluster 4 terdiri dari 121 desa.

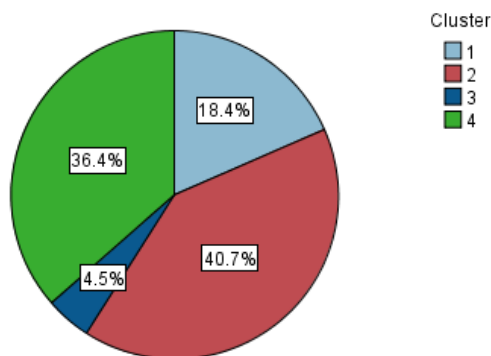
Hasil dan Pembahasan

Proses clustering desa Kabupaten Majalengka menggunakan algoritma *Two-Step Cluster* dilakukan proses pertama artinya, Bayesian Information Criterion (BIC) diperhitungkan dan kriteria keputusan optimal digunakan untuk menentukan jumlah cluster yang optimal. Berikut hasil perhitungan menggunakan indikator BIC. Selain menentukan estimasi jumlah cluster, juga ditentukan penentuan kriteria, perubahan BIC, rasio perubahan BIC dan rasio ukuran jarak.

Number of Clusters	Schwarz's Bayesian Criterion (BIC)	BIC Change ^a	Ratio of BIC Changes ^b	Ratio of Distance Measures ^c
1	1206.174			
2	972.305	-233.869	1.000	1.411
3	823.466	-148.839	.636	1.773
4	764.835	-58.630	.251	1.856
5	760.003	-4.832	.021	1.253
6	767.856	7.853	-.034	1.140
7	781.886	14.030	-.060	1.334
8	806.941	25.055	-.107	1.034
9	833.086	26.145	-.112	1.032
10	860.218	27.132	-.116	1.195
11	892.400	32.182	-.138	1.219
12	929.238	36.838	-.158	1.126
13	968.456	39.218	-.168	1.144
14	1010.052	41.596	-.178	1.138
15	1053.641	43.589	-.186	1.166

Gambar 4. Hasil Rasio

Hasil diatas merupakan hasil estimasi cluster yang terbentuk dengan menggunakan indikator BIC. Terdapat 15 indikator dengan detail perubahan jarak dan perhitungan jarak. Penentuan jumlah cluster pada BIC terlihat pada jarak rasio terbesar. Dari data penduduk di Kabupaten Majalengka yang terdapat 332 desa tersebut jumlah cluster yang memiliki rasio jarak terbesar adalah 4 cluster. Pada cluster 4 memiliki Ratio of distance measure tertinggi yaitu 1.856 dengan nilai BIC 764.835. Pada penelitian ini terbentuk jumlah cluster optimal yaitu empat cluster. Seperti yang terlihat pada Gambar 4.5.



Gambar 5. Kontribusi Margin

Pada margin kontribusi yang ditunjukkan pada Gambar 4.5, menunjukkan bahwa, Cluster 1 yang ditandai biru muda dengan 18,4% desa, cluster 2 ditandai merah dengan 40,7% desa, cluster 3 ditandai biru tua dengan 4,5%, dan cluster 4 yang ditandai dengan warna hijau dengan jumlah desa sebanyak 36,4%. Pada Tabel 4.7 menunjukkan hasil sebaran atau jumlah desa dari masing-masing cluster.

Tabel 4. 7 Distribusi cluster

Sebaran hasil cluster yang terbentuk dari 332 desa di Kabupaten Majalengka menunjukkan 4 cluster. Dengan cluster 1 memiliki 61 desa atau 18,4%, dan cluster 2 memiliki desa paling banyak yaitu 135 desa atau 40,7%, cluster 3 dengan jumlah desa paling sedikit dibandingkan cluster 1 dan cluster 2 yaitu 15 desa atau 4,5% dan cluster terakhir atau cluster 4 memiliki jumlah desa sebanyak 121 desa atau 36,4%. Seluruh desa di kecamatan tersebut akan mendapatkan hasil pengelompokan keanggotaan desa di Kabupaten Majalengka berdasarkan proses algoritma pengelompokan desa di Kabupaten Majalengka menggunakan metode Two-Step Cluster. Tabel 4.8 terdapat daftar keanggotaan untuk setiap cluster. Berdasarkan hasil cluster pada tabel dibawah ini, setiap desa dikelompokkan menurut jumlah laki-laki, jumlah perempuan, jumlah penduduk, jumlah kepemilikan kartu keluarga, dan jumlah kepadatan penduduk (jiwa/km²).

IV. KESIMPULAN

Berdasarkan hasil pengolahan dan analisis data, peneliti menarik kesimpulan yaitu Perhitungan pada nilai *cluster* dengan data kepadatan penduduk di Kabupaten Majalengka ini berdasarkan jumlah laki-laki, jumlah perempuan, jumlah penduduk, jumlah kartu keluarga dan kepadatan penduduk sehingga menghasilkan *cluster* optimal sebanyak 4 *cluster* dengan nilai *Ratio Of distance measure* sebesar 1.856 dengan nilai *Ratio of BIC Changes* sebesar 251.

Pada 4 cluster tersebut diperoleh pada cluster 1 sebanyak 18,4%, cluster 2 memiliki jumlah desa paling banyak yaitu 40,7%, cluster 3 hanya 4,5% dan cluster 4 memiliki 36,4% desa. Sehingga diperoleh cluster dengan jumlah penduduk paling padat pada cluster ke 3 dengan jumlah desa paling sedikit yaitu hanya 4,5% dengan jumlah penduduk paling padat yaitu 8.135 Jiwa/Km².

V. REFERENCES

- [1] r. M. N. Halim, "penerapan network attached storage (nas) berbasis raspberry pi di lp3sdm azra palembang," *J. Teknol. Inf. Dan ilmu komput.*, vol. 6, no. 3, p. 309, 2019, doi: 10.25126/jtiik.2019631416.
- [2] m. Jannah, b. L. Basyah, and r. A. Riyadi, "rancang bangun network attached storage (nas) pada raspberry pi untuk penyimpanan data terpusat berbasis wlan," *J. Ilm. Fifo*, vol. 7, no. 2, p. 222, 2015, doi: 10.22441/fifo.v7i2.1257.
- [3] k. I. Santoso and m. A. Muin, "implementasi network attached storage (nas) menggunakan nas4free untuk media backup file," *sci. J. Informatics*, vol. 2, no. 2, p. 123, 2016, doi: 10.15294/sji.v2i2.5078.
- [4] c. E. Suharyanto and a. Maulana, "perancangan network attached storage (nas) menggunakan raspberry pi untuk usaha mikro kecil dan menengah (umkm)," *jitk (jurnal ilmu pengetah. Dan teknol. Komputer)*, vol. 5, no. 2, pp. 271–278, 2020, doi: 10.33480/jitk.v5i2.1215.
- [5] a. Tanzeh and s. Arikunto, "metode penelitian metode penelitian," *metod. Penelit.*, pp. 22–34, 2014.
- [6] t. Sanjaya and d. Setiyadi, "network development life cycle (ndlc) dalam perancangan jaringan komputer pada rumah shalom mahanaim," *mhs. Bina insa.*, vol. 4, no. 1, pp. 1–10, 2019, [online]. Available: <http://ejournal-binainsani.ac.id/>.
- [7] w. Budiaji, "skala pengukuran dan jumlah respon skala likert (the measurement scale and the number of responses in likert scale)," *ilmu pertan. Dan perikan.*, vol. 2, no. 2, pp. 127–133, 2013, [online]. Available: <http://umbidharma.org/jipp>.
- [8] m. T. Kurniawan, a. Nurfaraj, o. Dwi, and u. Yunan, "desain topologi jaringan kabel nirkabel pdii-lipi dengan cisco three-layered hierarchical menggunakan ndlc," *elkomika j. Tek. Energi elektr. Tek. Telekomun. Tek. Elektron.*, vol. 4, no. 1, p. 47, 2018, doi: 10.26760/elkomika.v4i1.47.
- [9] n. Suarna, y. A. Wijaya, mulyawan, t. Hartati, and t. Suprapti, "comparison k-medoids algorithm and k-means algorithm for clustering fish cooking menu from fish dataset," *iop conf. Ser. Mater. Sci. Eng.*, vol. 1088, no. 1, p. 012034, 2021, doi: 10.1088/1757-899x/1088/1/012034.
- [10] s. Turangga and y. A. W, "analisis internet menggunakan parameter quality of service pada alfamart tuparev 70," vol. 6, no. 1, pp. 392–398, 2022.
- [11] t. Hartati and y. A. Wijaya, "analisis data lalu lintas jaringan di kantor cangehgar cyber operation center menggunakan algoritma k-means network traffic data analysis at cangehgar cyber operation center office using k-means algorithm," vol. 7, no. 1, pp. 75–84, 2022.
- [12] h. Putri, a. I. Purnamasari, a. R. Dikananda, o. Nurdiawan, and s. Anwar, "penerima manfaat bantuan non tunai kartu keluarga sejahtera menggunakan metode naïve bayes dan knn," *build. Informatics, technol. Sci.*, vol. 3, no. 3, pp. 331–337, 2021, doi: 10.47065/bits.v3i3.1093.

- [13] a. Z. Zami, o. Nurdiawan, and g. Dwilestari, "klasifikasi kondisi gizi bayi bawah lima tahun pada posyandu melati dengan menggunakan algoritma decision tree," *j. Sist. Komput. Dan inform.*, vol. 3, pp. 305–310, 2022, doi: 10.30865/json.v3i3.3892.
- [14] e. W. Ramadhona, t. Prasetya, and a. I. Purnamasari, "game edukasi ' nihongo kurabu ' belajar bahasa menggunakan unity 2d berbasis android," *inf. Manag. Educ. Prof.*, vol. 6, no. 1, pp. 71–80, 2022.
- [15] a. Z. Zami, o. Nurdiawan, and g. Dwilestari, "klasifikasi kondisi gizi bayi bawah lima tahun pada posyandu melati dengan menggunakan algoritma decision tree," *j. Sist. Komput. Dan inform.*, vol. 3, pp. 305–310, 2022, doi: 10.30865/json.v3i3.3892.
- [16] h. Putri, a. I. Purnamasari, a. R. Dikananda, o. Nurdiawan, and s. Anwar, "penerima manfaat bantuan non tunai kartu keluarga sejahtera menggunakan metode naïve bayes dan knn," *build. Informatics, technol. Sci.*, vol. 3, no. 3, pp. 331–337, 2021, doi: 10.47065/bits.v3i3.1093.
- [17] d. A. K. Irfan nurdiyanto, odi nurdiawan, ning rahaningsih, ade irfma purnamasari, "penentuan keputusan pemberian pinjaman kredit menggunakan algoritma c.45," *j. Data sci. Dan inform.*, vol. 1, no. 1, pp. 16–20, 2021.
- [18] a. Faqih, o. Nurdiawan, and a. Setiawan, "ethnomathematics : utilization of crock , ladle , and chopping board for learning material of geometry at the elementary school," vol. 4, no. 1, pp. 46–55, 2021.
- [19] o. Nurdiawan, f. A. Pratama, d. A. Kurnia, kaslani, and n. Rahaningsih, "optimization of traveling salesman problem on scheduling tour packages using genetic algorithms," *j. Phys. Conf. Ser.*, vol. 1477, no. 5, 2020, doi: 10.1088/1742-6596/1477/5/052037.
- [20] f. Arie pratama, k. Kaslani, o. Nurdiawan, n. Rahaningsih, and n. Nurhadiansyah, "learning innovation using the zahir application in improving understanding of accounting materials," *j. Phys. Conf. Ser.*, vol. 1477, no. 3, pp. 0–6, 2020, doi: 10.1088/1742-6596/1477/3/032018.