

Optimasi Algoritma *K-Means* Menggunakan Metode *Elbow* Pada Data Penerima Program Keluarga Harapan (PKH)

Castaka Agus Sugianto¹, Keny Kirana Wanaziana²

^{1,2}Program Studi Teknik Informatika-Politeknik TEDC, Jl. Pesantren km 2 Cibabat-Cimahi Utara, Cimahi, Indonesia

e-mail: castaka@poltektedc.ac.id¹, kenykwz@gmail.com²

INFORMASI ARTIKEL

Sejarah Artikel:

Diterima Redaksi : 2 Maret 2024

Revisi Akhir : 15 Mei 2024

Diterbitkan Online : 30 Mei 2024

Kata Kunci:

PKH, Optimasi, Algoritma *K-Means*, *Elbow*, Cimahi Utara

Korespondensi :

Telepon / Hp : +62 85352086591

E-mail : castaka@poltektedc.ac.id

A B S T R A K

Program Keluarga Harapan (PKH) merupakan program pemberian bantuan sosial bersyarat kepada keluarga miskin yang ditujukan kepada keluarga penerima manfaat PKH. Melalui PKH, keluarga miskin didorong untuk memiliki akses dan memanfaatkan pelayanan sosial dasar di bidang kesehatan, pendidikan, pangan dan gizi, perawatan, serta pendampingan. Pada kelurahan Cibabat dan kelurahan Cipageran kecamatan Cimahi Utara terdapat 387 data penerima bantuan PKH pada tahun 2022. Namun belum adanya pengolahan data penerima bantuan PKH tersebut sehingga dalam melakukan pendampingan penerima bantuan PKH belum mendapatkan pengelompokan yang sesuai dengan riwayat pendidikan. Berdasarkan latar belakang masalah yang telah dijabarkan, penulis tertarik untuk melakukan Clustering menggunakan Algoritma *K-means* pada data penerima bantuan PKH. Berdasarkan pengujian metode *elbow* pada algoritma *k-means* didapat nilai *k* yang optimal adalah *k*=3. Pengelompokan dataset yang digunakan menjadi 3 kelompok cluster, diantaranya cluster_0 sebanyak 257 data, cluster_1 sebanyak 75 data, dan cluster_2 sebanyak 55 data. Pada cluster_0 di dominasi oleh peserta lulusan SD sebanyak 173 data, untuk cluster_1 di dominasi oleh peserta tidak sekolah sebanyak 40 data, dan untuk cluster_2 di dominasi peserta tidak sekolah sebanyak 48 data. Pada cluster tersebut didapatkan nilai performa berdasarkan rata-rata avg. within centroid distance_cluster_0 adalah 6.720, avg. within centroid distance_cluster_1 adalah 14.373, avg. within centroid distance_cluster_2 adalah 8.496 dan Davies Bouildin Index adalah 0.816. Hasil penelitian ini diharapkan menjadi acuan bagi pengurus sekretariat PPKH dalam melaksanakan pendampingan masyarakat penerima Program Keluarga Harapan..

1. PENDAHULUAN

Peraturan Menteri Sosial Republik Indonesia nomor 1 Tahun 2018 tentang Program Keluarga Harapan dijelaskan bahwa program pemberian bantuan sosial bersyarat kepada keluarga dan/atau seseorang miskin dan rentan yang terdaftar dalam data terpadu program penanganan fakir miskin, diolah oleh Pusat Data dan Informasi Kesejahteraan Sosial dan ditetapkan sebagai keluarga penerima manfaat PKH [1]. [Program Keluarga Harapan](#) (PKH) adalah program pemberian bantuan sosial bersyarat kepada Keluarga Miskin (KM) yang ditujukan kepada keluarga penerima manfaat PKH. Sedikitnya 4.751 Rumah Tangga Sasaran Miskin (RTSM) di Kota Cimahi menjadi penerima bantuan PKH. Masyarakat yang mendapatkan bantuan dari PKH tidak hanya menerima bantuan berupa uang tunai tetapi kegiatan sehari-hari masyarakat juga dipantau jika ada yang menyimpang atau melanggar dan tidak memenuhi kewajiban akan diberikan sanksi [2]. Berdasarkan hasil observasi yang dilakukan oleh penulis di kelurahan Cibabat dan kelurahan Cipageran kecamatan Cimahi Utara terdapat 387 data penerima bantuan PKH pada tahun 2022. Namun belum adanya pengolahan data penerima bantuan PKH tersebut sehingga dalam melakukan pendampingan penerima bantuan PKH belum mendapatkan pengelompokan yang sesuai dengan riwayat pendidikan karena

pendidikan memiliki pengaruh dalam pola pikir masyarakat untuk menghadapi permasalahan dan penerimaan informasi [3]. Maka dari itu perlu dilakukan pengelompokan peserta penerima bantuan PKH pada hal ini untuk menyelesaikan permasalahan di atas di perlukan metode yang cepat dalam mengelompokan data salah satunya *data mining*. Dengan definisi *data mining*, ada beberapa teknik dan sifat analisa yang dapat digolongkan dalam *data mining* yaitu klasifikasi, *clustering*, *association rule*, *regression* [4].

Pada kali ini peneliti menggunakan teknik *Clustering* untuk mengelompokan data penerima bantuan PKH pada kecamatan Cimahi Utara. *Clustering* merupakan teknik untuk mengelompokkan data dan membentuk kelas objek-objek yang memiliki kemiripan. Dalam penelitian ini peneliti juga menggunakan algoritma *K-means*. Kelebihan dari algoritma *K-means clustering* adalah metode partisi yang paling efisien dan telah digunakan dalam berbagai aplikasi karena kesederhanaan dan dapat mencapai konvergensi yang cepat [5]. Algoritma *K-means* lebih baik daripada algoritma *K-Medoids* [6][7]. Dari segi waktu, algoritma *K-means* dapat mengkluster lebih cepat jika dibandingkan dengan algoritma [8]. Algoritma *K-means* sering digunakan oleh beberapa peneliti salah satunya [9] menerapkan *K-means* untuk mengelompokan penerima bantuan social menjadi 2 cluster. Metode ini

juga pernah digunakan untuk mengelompokkan bantuan bedah rumah dan membaginya ke dalam 3 *cluster*[10]. Penelitian lainnya juga yang menggunakan algoritma *k-means* yang menghasilkan 3 *cluster*, untuk mengelompokkan rumah tangga sangat miskin yang mendapatkan bantuan PKH[11]. Namun, algoritma *K-Means* memiliki kelemahan, terutama ketika menentukan jumlah *cluster* dan memilih *centroid* awal, yang dapat mempengaruhi hasil *clustering*[12][13]. Banyak metode yang bisa digunakan untuk mengoptimalkan jumlah *cluster* pada metode *clustering* diantaranya metode *elbow*, *Silhouette* dan *Particle Swarm Optimization (PSO)*[14][15] [16]. Metode *elbow* merupakan metode yang dapat diterapkan pada *clustering* untuk menentukan jumlah *cluster* yang optimal yang akan dibentuk[17]. Tujuan dari penelitian ini adalah untuk menentukan jumlah *cluster* yang optimal untuk pengklasteran penerima Program Keluarga Harapan (PKH).

2. METODE PENELITIAN

Dalam penelitian ini Ada 8 (delapan) tahapan yang akan dilakukan diantaranya:

A. Menentukan Masalah

Penentuan masalah yang peneliti lakukan merupakan proses awal dalam melakukan penelitian. Tahapan menentukan masalah sudah dijelaskan dalam latar belakang mengenai pengelompokan data penerima bantuan PKH menggunakan metode *clustering K-means*.

B. Menentukan Tujuan dan Ruang Lingkup

Menentukan tujuan dan ruang lingkup pada penulisan adalah tentang apa saja yang akan di bahas pada penulisan ini serta batasan-batasannya yang dibuat oleh peneliti agar peneliti fokus dalam pada masalah yang sudah diangkat.

C. Mencari Literature

Mencari literature atau referensi dilakukan untuk mencari berbagai sumber yang berkaitan tentang bantuan PKH, *data mining*, *clustering*, *RapidMiner*, dataset, algoritma *K-means*, dan metode *elbow*. Selain itu peneliti juga menggunakan metode studi pustaka yang merupakan metode mengumpulkan dan mencari sumber dari buku, internet, jurnal. Data-data yang dikumpulkan oleh penulis dari metode studi pustaka adalah:

- 1) Konsep *Data Mining*
- 2) Konsep *RapidMiner*
- 3) Konsep *Clustering*
- 4) Konsep *K-means*
- 5) Konsep metode *elbow*

D. Pengumpulan Data

Penelitian ini menggunakan beberapa sumber data diantaranya yaitu:

- 1) Data Primer
Data primer adalah data yang didapatkan penulis secara langsung dari sumber data

tersebut yang berkaitan dengan penelitian yang sedang dilakukan. Pada penelitian ini data yang didapat langsung dari Sekertariat Unit Pelaksana Program Keluarga Harapan berupa data peserta PKH tahun 2022.

2) Data Sekunder

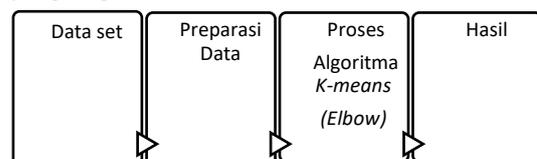
Data sekunder adalah data yang diperoleh secara tidak langsung, yaitu dari pihak kedua yang mengolah data untuk orang lain. Pada penelitian ini data sekunder yang didapat yaitu dari journal, artikel, buku yang berkaitan dengan penelitian.

3) Observasi

Observasi merupakan metode yang digunakan untuk memperoleh data dengan cara melakukan pertemuan tatap muka, pencarian data dan informasi, serta wawancara atau dengan alat perantara. Dalam penelitian ini melakukan observasi di kantor Sekertariat Unit Pelaksana Program Keluarga Harapan.

E. Metode Yang Digunakan

Tahap ini akan menjelaskan tentang proses dalam pengolahan data penulisan. Pada penelitian kali ini yang tergambar melalui Gambar 1.



Gambar 1. Proses Pengolahan Data

1. Dataset Masyarakat Penerima Bantuan PKH

Dataset yang digunakan dalam penelitian ini adalah data dari peserta penerima bantuan PKH kelurahan Cibabat dan Cipageran dengan persetujuan dari Dinas Sosial Kota Cimahi. Data yang didapatkan berupa data penerima bantuan PKH yang berjumlah 387 data dan data yang diambil dari periode tahun 2022.

2. Preparasi Data

Data yang sudah didapatkan ditahap ini akan di preparasi atau *cleaning data* untuk menghilangkan data yang kurang penting dalam proses pengolahan data nantinya. Gambar 2. Merupakan data yang sudah dipreparasi.

NO	NAMA_PENERIMA	ALAMAT	PEKERJAAN	JUMLAH ANAK	PDDKN
1	P0	RW 01	1	2	0
2	P1	RW 01	0	10	0
3	P2	RW 01	0	2	2
4	P3	RW 01	1	2	2
5	P4	RW 01	1	2	3
6	P5	RW 01	1	2	0
7	P6	RW 01	1	1	0
8	P7	RW 01	1	2	1
9	P8	RW 01	0	4	0
10	P9	RW 01	1	4	1
11	P10	RW 01	0	2	2
12	P11	RW 01	0	4	1
13	P12	RW 01	1	1	1
14	P13	RW 01	0	3	0
15	P14	RW 01	0	2	0
16	P15	RW 01	0	4	3
17	P16	RW 01	1	3	1
18	P17	RW 06	0	3	0
19	P18	RW 06	1	4	0
.
.
387	P386	RW 018	1	3	3

Gambar 2. Dataset Penerima Bantuan PKH

3. Proses Algoritma

Penelitian kali ini menggunakan aplikasi *rapidminer*. Pada aplikasi *rapidminer* penulis akan melakukan proses *cluster* data peserta bantuan PKH, algoritma yang dipakai penulis adalah algoritma *K-means* dengan metode optimasi *Elbow*.

4. Hasil

Hasil dari penelitian ini adalah *cluster* dari data pendidikan peserta PKH di daerah Kelurahan Cibabat dan Cipageran Kecamatan Cimahi Utara. Dalam penelitian ini juga akan di hasilkan pengujian algoritma *K-means* terhadap *clustering* data peserta penerima bantuan PKH di Kelurahan Cipageran dan Cibabat yang akan menghasilkan nilai k terbaik dan mendapatkan jumlah cluster terbaik pada pengelompokan data terkait.

F. Proses Dan Hasil

Proses pada penelitian ini penulis menggunakan metode *elbow*. Penulis menggunakan metode ini karena untuk mengetahui berapa cluster yang harus di hasilkan agar mendapatkan hasil cluster yang terbaik terhadap data penerima bantuan PKH.

G. Analisis

Proses analisis akan dilakukan penulis untuk mengamati dan menganalisa hasil yang didapat dari proses data mining menggunakan *software rapidminer*.

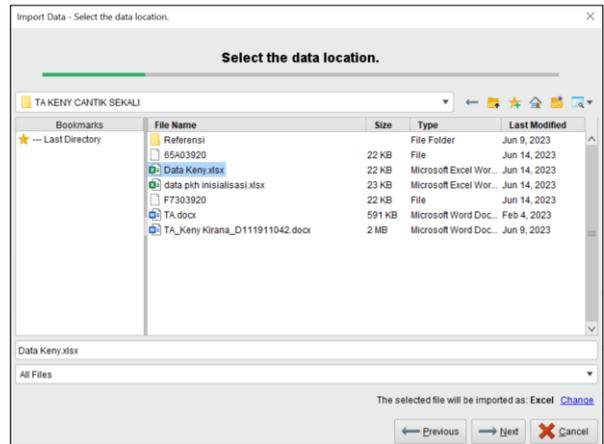
H. Penarikan Kesimpulan

Penarikan kesimpulan akan menjelaskan tentang hasil dari proses pengujian algoritma *k-means* dengan optimalisasi metode *elbow* dari data peserta penerima bantuan PKH di Kelurahan Cibabat dan Cipageran kecamatan Cimahi Utara.

3. HASIL DAN PEMBAHASAN

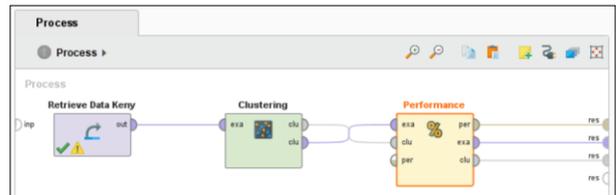
A. Menentukan Jumlah Cluster Terbaik

Tahap ini menggunakan metode *elbow* untuk menguji dan melihat hasil *cluster* terbaik dengan memasukan data yang sudah dipreparasi sebelumnya dengan cara *import* data pada *software rapidminer* dan menyeleksi data yang akan digunakan yaitu data dengan nama data *keny.xlsx* yang di beritanda warna biru. Dapat dilihat pada Gambar 3.



Gambar 3. Import Data

Proses pengujian ini menggunakan *software rapidminer* dan mengaplikasikan Algoritma *k-means* dengan memasukan data yang sudah dipreparasi sebelumnya. Proses pengujian tersebut dapat dilihat pada Gambar 4.



Gambar 4. Proses Memasukan Data Kedalam Rapidminer

Setelah beberapa proses di atas selanjutnya adalah mengatur parameter pada *K-means clustering*, penulis melakukan 9 kali percobaan untuk mendapatkan nilai *avg. within centroid distance* sebagai bahan untuk pengujian metode *elbow*.

Cluster Model

```

Cluster 0: 53 items
Cluster 1: 28 items
Cluster 2: 95 items
Cluster 3: 39 items
Cluster 4: 79 items
Cluster 5: 23 items
Cluster 6: 11 items
Cluster 7: 32 items
Cluster 8: 13 items
Cluster 9: 14 items
Total number of items: 387
        
```

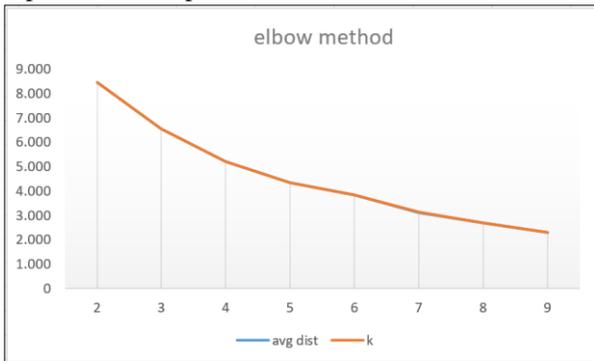
Avg. within centroid distance

Avg. within centroid distance: 2.300

Gambar 5. Hasil Pengujian parameter k-10

Setelah pengaturan parameter *k-means clustering* sudah selesai, maka selanjutnya penulis akan melakukan pengujian menggunakan metode *elbow* dari hasil ke 9 parameter diatas kita masukan kedalam *Microsoft excel* untuk melihat nilai manakah yang memiliki kurva paling siku yang dihasilkan metode *elbow* untuk menentukan nilai k yang tepat.

Gambar 6 menunjukkan kurva yang memiliki sudut yang tajam pada saat bernilai k=3 dan k=4. Dalam metode *elbow* ketika ada lebih dari satu nilai k yang membentuk siku maka dalam menentukan k terbaik dengan melihat seberapa besar nilai *avg. within centroid distance* pada kedua nilai k tersebut, untuk k=3, nilai *avg. within centroid distance* adalah 8.456, sedangkan untuk k=4, nilai *avg. within centroid distance* adalah 6.557. Oleh karena itu berdasarkan hasil ini, maka dapat disimpulkan bahwa k terbaik adalah k=3. Uji coba pengujian metode *elbow* juga dapat ditemukan pada Gambar 6.



Gambar 6. Hasil Pengujian Metode Elbow

B. Hasil Data Cluster K-Means

Dari proses penentuan *cluster* didapatkan *cluster* 3 dengan menggunakan metode *elbow* dengan nilai yang terbaik. Maka pengelompokan data pada penelitian ini menjai 3 kelompok yaitu *cluster_0* dengan jumlah 257 data, kelompok data *cluster_1* berjumlah 75 data sedangkan kelompok data *cluster_2* berjumlah 55 data dengan total 387 data dan nilai *avg. within distance cetrroid* adalah 8.456 sedangkan nilai *Davies Bouldin Index* 0.816. Pengelompokan data tersebut dapat dilihat pada Gambar 7 dan Gambar 8.

```

Cluster Model

Cluster 0: 257 items
Cluster 1: 75 items
Cluster 2: 55 items
Total number of items: 387
    
```

Gambar 7. Hasil Data Cluster K-Means

Dimana akurasi berdasarkan rata-rata *within centroid distance* dapat dilihat pada Gambar 8.

```

PerformanceVector

PerformanceVector:
Avg. within centroid distance: 8.456
Avg. within centroid distance_cluster_0: 6.720
Avg. within centroid distance_cluster_1: 14.373
Avg. within centroid distance_cluster_2: 8.496
Davies Bouldin: 0.816
    
```

Gambar 8. Rata-rata *centroid distance*

Berikut data peserta penerima bantuan PKH yang termasuk kedalam *cluster_0* terdiri dari 257 data dapat dilihat pada Tabel 1.

Tabel 1. Data yang Termasuk Kedalam *Cluster_0*

No	NAMA PENERIMA	CLUSTER	JK	ALAMAT	PEKERJAAN	JUMLAH ANAK	PENDIDIKAN
1	P0	cluster_0	P	1	Tidak Bekerja	2	SD
2	P1	cluster_0	P	1	Bekerja	10	SD
3	P2	cluster_0	P	1	Bekerja	2	SLT A
4	P3	cluster_0	P	1	Tidak Bekerja	2	SLT A
6	P5	cluster_0	P	1	Tidak Bekerja	2	SD
7	P6	cluster_0	P	1	Tidak Bekerja	1	SD
8	P7	cluster_0	P	1	Tidak Bekerja	2	SLTP
9	P8	cluster_0	P	1	Bekerja	4	SD
10	P9	cluster_0	P	1	Tidak Bekerja	4	SLTP
.
.
.
257	P311	cluster_0	L	8	Tidak Bekerja	3	SD

Berikut data peserta penerima bantuan PKH yang termasuk kedalam *cluster_1* yang terdiri dari 75 data dapat dilihat pada Tabel 2.

Tabel 2. Data yang Termasuk Kedalam *Cluster_1*

NO	NAMA PENERIMA	CLUSTER	JK	ALAMAT	PEKERJAAN	JUMLAH ANAK	PENDIDIKAN
1	P312	cluster_1	L	13	Bekerja	3	SD
2	P313	cluster_1	L	13	Bekerja	5	SD
3	P314	cluster_1	P	13	Bekerja	2	SD
4	P315	cluster_1	P	13	Tidak Bekerja	1	TS
5	P316	cluster_1	P	13	Bekerja	3	SD
.
.
.
75	P386	cluster_1	P	18	Tidak	3	TS

NO	NAMA PENERIMA	CLUSTER	JK	ALAMAT	PEKERJAAN	JUMLAH ANAK	PENDIDIKAN
Bekerja							

Berikut data peserta penerima bantuan PKH yang termasuk kedalam *cluster_1* yang terdiri dari 55 data dapat dilihat pada Tabel 3.

Tabel 3. Data yang Termasuk Kedalam *Cluster_2*

NO	NAMA PENERIMA	CLUSTER	JK	ALAMAT	PEKERJAAN	JUMLAH ANAK	PENDIDIKAN
1	P4	<i>cluster_2</i>	P	1	Tidak Bekerja	2	SM K
2	P15	<i>cluster_2</i>	P	1	Bekerja	4	SM K
3	P41	<i>cluster_2</i>	P	6	Tidak Bekerja	3	SLT P
4	P59	<i>cluster_2</i>	L	6	Tidak Bekerja	0	TS
5	P117	<i>cluster_2</i>	P	7	Tidak Bekerja	4	SLT A
6	P136	<i>cluster_2</i>	P	7	Tidak Bekerja	0	TS
7	P138	<i>cluster_2</i>	P	7	Tidak Bekerja	2	SLT P
8	P141	<i>cluster_2</i>	P	7	Tidak Bekerja	1	SLT A
9	P175	<i>cluster_2</i>	P	7	Tidak Bekerja	0	TS
10	P177	<i>cluster_2</i>	P	2	Tidak Bekerja	3	TS
.
.
.
55	P309	<i>cluster_2</i>	L	8	Tidak Bekerja	3	TS

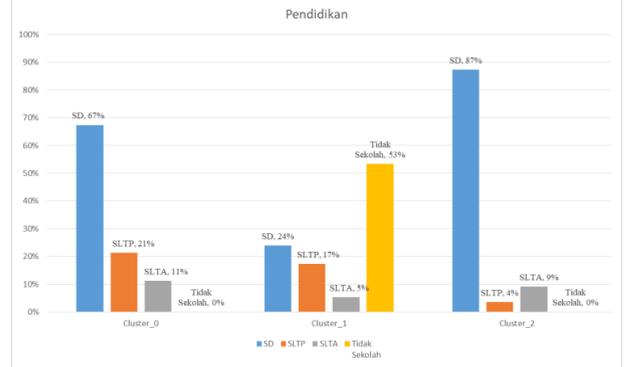
C. Analisis Hasil Algoritma K-means

Hasil pengujian yang sudah dilakukan oleh penulis menggunakan metode *elbow* pada algoritma *k-means* menghasilkan 3 *cluster*. Dari 387 peserta penerima bantuan PKH menghasilkan pengelompokan yang didominasi sesuai kategori pendidikan yaitu dengan Riwayat Pendidikan SD dan Tidak Sekolah.

D. Pengelompokan Data Penerima Bantuan PKH

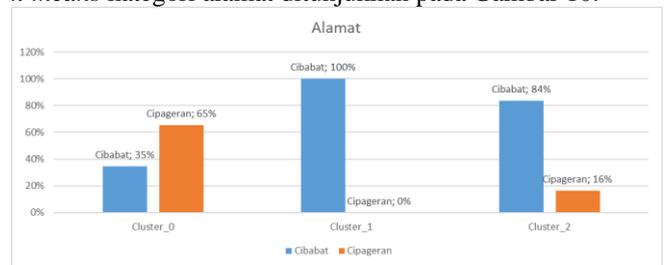
Dari hasil pengujian menggunakan metode *elbow* didapatkan hasil *cluster* terbaik yaitu k=3 berdasarkan kategori pendidikan. Dengan itu untuk pengelompokan *dataset* yang digunakan menjadi 3 kelompok *cluster* berdasarkan kategori pendidikan, diantaranya *cluster_0* sebanyak 257, *cluster_1* sebanyak 75 data, dan *cluster_2* sebanyak 55 data. Dilihat dari Tabel 4.1, 4.2 dan 4.3 hasil dari *cluster k-means* terdapat beberapa dominasi data dari setiap *clusternya*. Pada *cluster_0* didominasi oleh peserta lulusan SD sebanyak 173 data, SLTP sebanyak 55 data dan SLTA sebanyak 29 data, untuk *cluster_1* didominasi oleh

peserta tidak sekolah sebanyak 40, SD sebanyak 18 data, SLTP sebanyak 13 data, dan SLTA 4 data untuk *cluster_2* didominasi peserta tidak sekolah sebanyak 48 data, SLTA sebanyak 5 data, SLTP sebanyak 2 data. Hal ini dapat menunjukkan bahwa penerima bantuan PKH di Kecamatan Cimahi Utara yaitu lulusan SD dan yang tidak bersekolah. Grafik hasil pengujian menggunakan metode *elbow* pada algoritma *k-means* kategori pendidikan ditunjukkan pada Gambar 9.



Gambar 9. Grafik Kategori Pendidikan

Berdasarkan dari alamat *cluster_0* didominasi oleh kelurahan Cipageran dengan jumlah 168 data diantaranya RW 1 sebanyak 16 data, RW 6 sebanyak 81 data dan RW 7 sebanyak 71 data, sedangkan kelurahan Cibabat yaitu 89 data diantaranya RW 2 sebanyak 14 data, RW 5 sebanyak 31 data dan RW 8 sebanyak 44 data, lalu *cluster_1* didominasi dengan kelurahan Cibabat dengan jumlah 75 data diantaranya RW 13 sebanyak 24 data dan RW 18 sebanyak 51 data, sedangkan *cluster_2* didominasi dengan kelurahan Cibabat dengan jumlah 46 data diantaranya RW 2 sebanyak 14 data, RW 5 sebanyak 13 data, dan RW 8 sebanyak 19 data, sedangkan untuk kelurahan Cipageran yaitu 9 data diantaranya RW 1 sebanyak 2 data, RW 6 sebanyak 2 data, dan RW 7 sebanyak 5 data. Hasil *cluster* berdasarkan kategori alamat menggunakan algoritma *k-means* menghasilkan daerah kelurahan yang menerima bantuan PKH yaitu didominasi oleh Kelurahan Cipageran Kecamatan Cimahi Utara dengan jumlah data 168 data. Grafik hasil pengujian menggunakan metode *elbow* pada algoritma *k-means* kategori alamat ditunjukkan pada Gambar 10.



Gambar 10. Grafik Kategori Alamat

E. Upaya Untuk Mengurangi Angka Kemiskinan

Hasil analisis mengenai pengelompokan data penerima bantuan PKH yang menggunakan algoritma *k-means* maka penulis dapat menyimpulkan beberapa saran untuk menekan angka kemiskinan di wilayah

kelurahan Cibabat dan Cipageran, maka dari itu untuk pengembangan lebih lanjut maka diperlukan beberapa saran sebagai berikut:

1. Perlunya pendampingan dilakukan kepada masyarakat untuk memberikan pemahaman mengenai pentingnya pendidikan. Hal ini disebabkan karena pendidikan yang baik dapat membantu mengurangi kemiskinan jangka panjang dan dengan memperoleh pendidikan yang layak untuk memiliki peluang yang lebih baik mendapatkan pekerjaan yang layak dan meningkatkan pendapatan.
2. Hasil pengelompokan dari data penerima bantuan PKH didominasi oleh peserta yang riwayat pendidikannya adalah tingkat Sekolah Dasar dan tidak sekolah, maka pada saat pendampingan pendamping dari masing-masing kelurahan harus memberikan pelatihan keterampilan yang relevan, mencari peluang ekonomi dan memberikan informasi tentang bantuan modal usaha, pelatihan kewirausahaan atau memberikan saran mengenai jenis usaha yang sesuai kemampuan.

4. KESIMPULAN

Berdasarkan hasil pengujian yang telah dilakukan, maka dapat ditarik kesimpulan sebagai berikut:

1. Pengujian menggunakan metode *elbow* pada algoritma *k-means* didapatkan hasil 3 *cluster*, dimana *cluster_0* sebanyak 257 data, *cluster_1* sebanyak 75 data, dan *cluster_2* sebanyak 55 data dengan total 387 data. Hasil *cluster* berdasarkan kategori riwayat pendidikan menggunakan algoritma *k-means* menghasilkan riwayat pendidikan pada peserta bantuan PKH di Kelurahan Cibabat dan Cipageran Kecamatan Cimahi Utara yaitu didominasi oleh lulusan SD dan tidak sekolah. Berdasarkan kategori alamat menggunakan algoritma *k-means* menghasilkan daerah kelurahan yang menerima bantuan PKH yaitu didominasi oleh Kelurahan Cipageran Kecamatan Cimahi Utara dengan jumlah data 168 data.
2. Berdasarkan pengujian yang telah dilakukan oleh penulis menggunakan metode *elbow* untuk mencari *cluster* terbaik dilihat dari kurva paling siku diantara 9 percobaan lain untuk menentukan *k* terbaiknya, *k* yang dihasilkan dari pengujian *elbow* yaitu *k*=3 maka dapat disimpulkan bahwa dalam metode *elbow* *k* terbaiknya yaitu pada *k*=3. Sedangkan untuk algoritma *k-means*, didapatkan nilai performa berdasarkan rata-rata *avg. within centroid distance_cluster_0* adalah 6.720, *avg. within centroid distance_cluster_1* adalah 14.373, *avg. within*

centroid distance_cluster_2 adalah 8.496 dan *Davies Bouildin Index* adalah 0.816.

DAFTAR PUSTAKA

- [1] Kementrian Sosial RI, "Peraturan Menteri Sosial Republik Indonesia Nomor 1 Tahun 2018 Tentang Program Keluarga Harapan Dengan Rahmat Tuhan Yang Maha Esa Menteri Sosial Republik INDONESIA, Menimbang," 2018.
- [2] Z. A. Fakrulloh, "Dukcapil Kemendagri rilis data penduduk semester 1 tahun 2022, naik 0,54% dalam waktu 6 bulan," *Direktorat Jendral Kependudukan dan Pencatatan Sipil Kementerian Dalam Negeri*. p. 1, 2022.
- [3] D. Nafkiyah, L. Rifatin, and M. R. Rozikin, "Analisis Cluster dalam Pengelompokan Kabupaten / Kota di Provinsi Jawa Timur berdasarkan Indikator Pendidikan," *J. Ilm. Mat. dan Pendidik. Mat.*, vol. 12, no. 1, pp. 1–16, 2022.
- [4] M. A. Muslim *et al.*, "Buku Data Mining," 2019.
- [5] F. Fajriani, "Persebaran Tingkat Kemiskinan di Sulawesi Selatan," in *Seminar Nasional Teknologi Informasi dan Komputer*, 2019, pp. 276–281.
- [6] C. A. Sugianto, A. H. Rahayu, and A. Gusman, "Algoritma K-Means untuk Pengelompokan Penyakit Pasien pada Puskesmas Cigugur Tengah," *J. Inf. Technol.*, vol. 2, no. 2, pp. 39–44, 2020, doi: 10.47292/joint.v2i2.30.
- [7] C. A. Sugianto and T. P. O. R. Bokings, "K-Means Algorithm For Clustering Poverty Data in Bangka Belitung Island Province," *J. Comput. Networks , Archit. High Perform. Comput.*, vol. 3, no. 1, pp. 58–67, 2021.
- [8] W. Widayani and H. Harliana, "Perbandingan Algoritma K-Means dan SFCM Pada Pengelompokan Rumah Tangga Miskin," *J. Sains dan Inform.*, vol. 6, no. 1, pp. 1–9, 2020, doi: 10.34128/jsi.v6i1.200.
- [9] L. G. Rady Putra and A. Anggrawan, "Pengelompokan Penerima Bantuan Sosial Masyarakat dengan Metode K-Means," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 21, no. 1, pp. 205–214, 2021, doi: 10.30812/matrik.v21i1.1554.
- [10] Y. Kusnadi and M. S. Putri, "Clustering Menggunakan Metode K-Means Untuk Menentukan Prioritas Penerima Bantuan Bedah Rumah (Studi Kasus : Desa Ciomas Bogor)," *J. Teknol. Inform. dan Komput.*, vol. 7, no. 1, pp. 17–24, 2021, doi: 10.37012/jtik.v7i1.498.
- [11] S. A. Azis, S. Defit, and Y. Yunus, "Klasterisasi Dana Bantuan Pada Program Keluarga Harapan (PKH)," *J. Inform. Ekon. Bisnis*, vol. 3, pp. 53–59, 2021, doi:

- 10.37034/infeb.v3i2.66.
- [12] A. Muqoddam, “Pengelompokan Produksi Tambak Garam Dengan Metode Cluster K-Means Dan Optimasi Cluster Menggunakan Elbow (Studi Kasus: Dinas Kelautan Kabupaten Bangkalan),” *Method. J. Tek. Inform. dan Sist. Inf.*, vol. 9, no. 1, pp. 45–49, 2023, doi: 10.46880/mtk.v9i1.1586.
 - [13] M. Orisa, “Optimasi Cluster pada Algoritma K-Means,” in *Prosiding SENIATI*, 2022, pp. 430–437, doi: 10.36040/seniati.v6i2.5034.
 - [14] Handrik, Kusri, and Kusnawi, “Optimasi Penentuan Sentroid Awal Pada K-Means Untuk Meningkatkan Hasil Evaluasi Davies Bouldin Index,” *J. Inform. Teknol. dan Sains*, vol. 6, no. 1, pp. 52–57, 2024.
 - [15] R. Saputra and I. Purnama, “optimizing k-means algorithm with elbow and silhouette methods for national exam score data clustering,” *J. Ilmu Komput. Ruru*, vol. 1, no. 1, pp. 1–6, 2024.
 - [16] A. R. Lashiyanti, I. R. Munthe, and F. A. Nasution, “Optimisasi Klasterisasi Nilai Ujian Nasional dengan Pendekatan Algoritma K-Means, Elbow, dan Silhouette,” *J. Ilmu Komput. dan Sist. Inf. (JIKOMSI)*, vol. 6, no. 1, pp. 14–20, 2023.
 - [17] N. A. Maori and E. Evanita, “Metode Elbow dalam Optimasi Jumlah Cluster pada K-Means Clustering,” *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 14, no. 2, pp. 277–288, 2023, doi: 10.24176/simet.v14i2.9630.