

Penggerombolan Desa di Kabupaten Poso Berdasarkan Sarana Prasarana dan Tenaga Kesehatan Menggunakan Metode *K-Prototype* dengan Algoritma Genetika

SEPTYANI¹, NUR'ENI², LILIES HANDAYANI³

¹Mahasiswa Prodi Statistika, Universitas Tadulako

^{2,3}Dosen Statistika, Universitas Tadulako

Email: reythakawuwungseptyani@gmail.com

ABSTRAK

Analisis gerombol merupakan salah satu teknik untuk mengelompokkan sekumpulan objek yang mirip dengan properti yang sama menjadi satu kelompok sehingga tidak mirip dengan objek di kelompok lainnya. Analisis gerombol umumnya diterapkan pada objek dengan tipe data numerik. Namun pada kenyataannya *clustering* juga menggunakan tipe data kategorikal. Penanganan *clustering* dengan data tipe campuran dapat dilakukan dengan menerapkan algoritma *k-prototype*, namun penentuan inisialisasi pusat cluster cenderung sensitif. Untuk menangani penentuan inisialisasi pusat cluster, dapat diterapkan suatu algoritma yaitu algoritma genetika. Penelitian ini membahas tentang sarana dan prasarana serta tenaga kesehatan di Kabupaten Poso dimana sarana prasarana dan tenaga hukum di kabupaten tersebut sudah memadai namun distribusinya tidak merata di beberapa daerah. Hasil penelitian ini menunjukkan bahwa pada algoritma *k-prototype* terdapat 8 cluster dengan pusat cluster yang dioptimasi menggunakan algoritma genetika yaitu 36, 7, 99, 49, 69, 104, 105, 110.

Kata Kunci: Analisis Gerombol, *K-prototype*, Algoritma Genetika, Infrastruktur, Tenaga Kesehatan.

ABSTRACT

Cluster analysis is a technique for grouping a set of similar objects into one group so that they are not similar to objects in other groups. Cluster analysis is generally applied to objects with numeric data type. But in reality clustering also uses categorical data types. Clustering handling with mixed-type data can be done by applying the *k-prototype* algorithm, but the determination of cluster center initialization tends to be sensitive. To handle the determination of the initialization of the cluster center, can be applied an algorithm that is genetic algorithm. This study discusses the facilities and infrastructure and health workers in Poso Regency where the infrastructure and legal personnel in the district are adequate but the distribution is not evenly distributed in several areas. The results of this study indicate that in the *k-prototype* algorithm there are 8 clusters with cluster centers optimized using the genetic algorithm, namely 36, 7, 99, 49, 69, 104, 105, 110.

Keywords: Cluster Analysis, *K-prototype*, Genetic Algorithm, Infrastructure, Health Workers.

1. PENDAHULUAN

Latar Belakang

Mengenai kesehatan provinsi Sulawesi Tengah pada tahun 2018, Sulawesi Tengah memiliki layanan dan penunjang kesehatan berupa sarana prasarana serta tenaga medis yang tersebar diseluruh daerah. Provinsi Sulawesi Tengah memiliki puskesmas yang berjumlah 206 unit, rumah bersalin yang berjumlah 13 unit, klinik yang berjumlah 104 unit, rumah sakit berjumlah 38 unit serta memiliki tenaga kesehatan terbanyak yaitu perawat, bidan, dan kesehatan masyarakat. Tenaga kesehatan minimal yang harus dimiliki oleh fasilitas kesehatan antara lain gizi, kesehatan lingkungan, kefarmasian dan keterampilan fisik masih sangat terbatas, serta jumlah dan jenis tenaga kesehatan belum merata pada setiap puskesmas. Belum seluruh puskesmas (45,8%) yang ada di provinsi Sulawesi Tengah memenuhi jumlah tenaga kesehatan minimal sesuai dengan

Permenkes 75 tahun 2014. Begitupun tenaga kesehatan pada rumah sakit, masih banyak rumah sakit belum memenuhi ketentuan standar jumlah dokter spesialis dan belum merata pada setiap rumah sakit di provinsi Sulawesi Tengah. Hal ini menjadi tantangan bagi Pemerintah Daerah (Provinsi, Kabupaten / Kota) dan pemerintah pusat untuk pemenuhan tenaga kesehatan (Dinkes, 2018).

Kabupaten Poso juga merupakan salah satu Kabupaten yang berada di Provinsi Sulawesi Tengah. Kabupaten Poso sendiri memiliki 19 kecamatan, 142 desa dan 28 kelurahan. Jumlah tenaga kesehatan Kabupaten Poso terdiri dari 61 tenaga kedokteran, 703 tenaga keperawatan, 483 tenaga kebidanan dan 94 tenaga kefarmasian dan sarana kesehatan yang terbilang memadai. Sarana prasarana dan tenaga kesehatan di kabupaten Poso memang sudah memadai akan tetapi penyebarannya belum merata. Berdasarkan informasi Publikasi Statistik Daerah Kabupaten Poso Tahun 2019 mengatakan bahwa pada tahun 2018 di kecamatan Lore Barat dan Lore Timur tidak ada satupun dokter yang dipekerjakan di kecamatan tersebut. Sehingga perlu dilakukan pengelompokkan wilayah untuk mengetahui wilayah mana saja yang perlu pembenahan dalam sarana kesehatan dan tenaga kesehatan (BPS, 2019).

Salah satu metode statistika yang umum digunakan untuk mengelompokkan objek adalah analisis gerombol (*cluster analysis*). Data sarana kesehatan dan tenaga kesehatan memiliki tipe kategorik dan numerik, sehingga perlu metode penggerombolan yang dapat diterapkan pada data campuran yaitu Algoritma *K-prototype*. *K-prototype* memiliki keunggulan dimana algoritmanya yang tidak terlalu kompleks dan mampu menangani data yang besar namun perlu menerapkan suatu teknik pengoptimasi yang dapat mengatasi masalah optimum lokal, salah satunya dengan menerapkan algoritma genetika. Algoritma genetika merupakan suatu alat optimasi yang dapat mengoptimalkan inisial pusat gerombol sehingga dapat diperoleh hasil penggerombolan yang global optimum. (Huang, 1997). Sehingga dari uraian tersebut penulis tertarik melakukan penelitian untuk menggerombolkan desa di kabupaten Poso dengan menerapkan metode *k-prototype* yang dioptimalkan dengan algoritma genetika untuk memperoleh solusi optimum global sehingga hasil penggerombolan menjadi lebih baik.

Rumusan Masalah

Berdasarkan latar belakang di atas maka rumusan masalah pada penelitian ini adalah bagaimana hasil penggerombolan desa di kabupaten Poso dengan menerapkan metode *k-prototype* yang dioptimalkan dengan algoritma genetika.

Tujuan Penelitian

Tujuan penelitian ini adalah mendapatkan hasil penggerombolan desa di kabupaten Poso dengan menerapkan metode *k-prototype* yang dioptimalkan dengan algoritma genetika.

Manfaat Penelitian

Berikut merupakan beberapa manfaat untuk penelitian ini yaitu:

1. Menambah wawasan akan metode yang digunakan, dimana merupakan salah satu metode penggerombolan yang dapat digunakan untuk menggerombolkan data dengan tipe campuran (numerik dan kategorik).
2. Memberikan informasi yang dapat dijadikan bahan pertimbangan bagi pihak pemerintah dalam melakukan pemerataan sarana prasarana dan tenaga kesehatan di setiap desa yang berada di kabupaten Poso.

Batasan Masalah

Berikut merupakan batasan masalah untuk penelitian ini yaitu:

1. Data yang digunakan merupakan data sarana prasarana dan tenaga kesehatan yang berada di setiap desa di kabupaten Poso tahun 2018.
2. Jumlah gerombol (*K*) yang digunakan sebanyak 8 gerombol.
3. Jumlah kromosom yang digunakan sebanyak 10.
4. Jumlah Generasi (iterasi) yang digunakan sebanyak 10 generasi.
5. Metode seleksi yang digunakan yaitu metode *roulette-wheel*.
6. Metode pindah silang (*crossover*) yang digunakan yaitu pindah silang satu titik (*one-point crossover*).
7. Metode mutasi yang digunakan yaitu *random*.

2. TINJAUAN PUSTAKA

Analisis Gerombol

Analisis gerombol merupakan salah satu teknik untuk mengelompokkan sekumpulan objek yang mirip dengan properti yang sama menjadi satu kelompok sehingga tidak mirip dengan objek di kelompok lainnya (Han & Kamber, 2006).

Dalam metode analisis gerombol diperlukan suatu ukuran jarak yang dijelaskan untuk setiap pasang objek yang akan dikelompokkan. Dalam analisis gerombol jarak yang biasa digunakan yaitu ukuran data numerik dan ukuran data kategorik.

Bentuk umum ukuran kesamaan dinyatakan sebagai berikut:

$$d(X_i, Z_l) = \sum_{j=1}^m \delta(x_{ij}, z_{lj})$$

Dimana:

$z_l = [z_{l1}, z_{l2}, \dots, z_{lm}]^T$: *prototype* untuk gerombol l .

Jarak *euclidean* merupakan ukuran kesamaan untuk atribut numerik yang ditunjukkan dalam persamaan berikut:

$$d(X_i, Z_l) = \left(\sum_{j=1}^{m_r} (x_{ij}^r - z_{lj}^r)^2 \right)^{1/2}$$

Dimana:

x_{ij}^r : nilai pada atribut numerik j

z_{lj}^r : rata-rata atau *prototype* atribut numerik ke j gerombol l

m_r : jumlah atribut numerik

Sedangkan ukuran kesamaan untuk data kategorik adalah

$$d(X_i, Z_l) = \gamma_l \sum_{j=l+1}^{m_c} \delta(x_{ij}^c, z_{lj}^c)$$

Dimana *simple matching similarity measure* untuk data kategorik adalah

$$\delta(x_{ij}^c, z_{lj}^c) = \begin{cases} 0 & (x_{ij}^c = z_{lj}^c) \\ 1 & (x_{ij}^c \neq z_{lj}^c) \end{cases}$$

Dimana:

γ_l : bobot untuk atribut kategorik pada gerombol l yang nilainya merupakan nilai standar deviasi untuk atribut numerik pada masing-masing gerombol

x_{ij}^c : nilai atribut kategorik

z_{lj}^c : modus atribut ke j gerombol l

m_c : jumlah atribut kategorik

Untuk meningkatkan kemiripan objek dalam gerombol dengan atribut kategorik sehingga hasil penggerombolan menjadi lebih baik maka dilakukan modifikasi *simple matching similarity measure*. Maka ukuran kesamaan untuk dataset yang memiliki atribut numerik dan atribut kategorik adalah

$$d(X_i, Z_l) = \left(\sum_{j=1}^{m_r} (x_{ij}^r - z_{lj}^r)^2 + \gamma_l \sum_{j=l+1}^{m_c} \delta(x_{ij}^c, z_{lj}^c) \right)^{1/2}$$

Dimana:

x_{ij}^r : nilai pada atribut numerik j

z_{lj}^r : rata-rata atau *prototype* atribut numerik ke j gerombol l

m_r : jumlah atribut numerik

γ_l : bobot untuk atribut kategorik pada gerombol l yang nilainya merupakan nilai standar deviasi untuk atribut numerik pada masing-masing gerombol

x_{ij}^c : nilai atribut kategorik

z_{ij}^c : modus atribut ke j gerombol l

m_c : jumlah atribut kategorik (Ahmad & Lipika, 2007).

Metode Elbow

Metode *elbow* merupakan suatu metode yang digunakan untuk menghasilkan informasi dalam menentukan jumlah gerombol terbaik dengan cara melihat presentase hasil perbandingan antara jumlah gerombol yang akan membentuk siku pada suatu titik. Hasil presentase yang berbeda dari setiap nilai gerombol dapat ditunjukkan menggunakan grafik sebagai sumber informasinya. Jika nilai gerombol pertama dengan nilai gerombol kedua memberikan sudut pada grafik atau nilainya mengalami penurunan paling besar maka nilai gerombol tersebut yang terbaik (Madhulatha, 2012).

Algoritma K-prototype

Algoritma *k-prototype* merupakan suatu metode gerombol yang dapat melakukan penggerombolan pada data dengan tipe campuran (numerik dan kategorik). Algoritma ini merupakan hasil pengembangan atau kombinasi dari *k-mean* dan *k-modes* untuk menangani penggerombolan pada data dengan atribut bertipe campuran (numerik dan kategorik). Pengembangan yang mendasar terdapat pada pengukuran kesamaan (*similarity measure*) antara objek dan *centroid (prototype)*-nya (Huang, 1998). Dilakukan beberapa proses pada penggerombolan dengan *k-prototype* yaitu:

- a. Menentukan jumlah gerombol (k) yang akan dibentuk.
- b. Menentukan inisial pada sejumlah gerombol (k) yang sudah ditentukan yaitu Z_1, Z_2, \dots, Z_k sebagai pusat cluster di masing-masing gerombol.
- c. Melakukan perhitungan jarak seluruh data observasi pada dataset terhadap inisial gerombol awal.
- d. Melakukan alokasi semua observasi ke dalam gerombol yang memiliki jarak *prototype* terdekat dengan objek yang diukur.
- e. Melakukan perhitungan titik pusat gerombol yang baru setelah semua objek dialokasikan. Kemudian merelokasikan semua data observasi pada dataset terhadap *prototype* yang baru.
- f. Apabila titik pusat gerombol tidak berubah atau sudah konvergen maka proses algoritma berhenti. Akan tetapi apabila titik pusat masih berubah-ubah secara signifikan maka proses kembali ke tahap (b) sampai dengan (e) hingga iterasi maksimum tercapai atau dengan kata lain sudah tidak ada lagi perpindahan objek.
- g. Melakukan proses penggerombolan algoritma ini pada ukuran k yang lain.

Algoritma Genetika

Algoritma genetika merupakan suatu algoritma pencarian berbasis pada proses seleksi alam dan genetika. Salah satu algoritma yang sangat tepat untuk digunakan dalam menyelesaikan masalah optimasi kompleks yang sulit dilakukan oleh metode konvensional yaitu algoritma genetika (Desiani & Arharmi, 2006).

1. Membangkitkan Populasi Awal

Membangkitkan populasi awal merupakan suatu proses untuk membangkitkan sejumlah individu secara acak atau melalui prosedur tertentu. Dimana terlebih dahulu harus menentukan jumlah individu dalam populasi tersebut sebelum melakukan pembangkitan populasi awal. Misalnya jumlah individu tersebut N . Setelah itu, baru kita membangkitkan populasi awal yang mempunyai N individu secara *random*. Syarat-syarat yang harus dipenuhi untuk menunjukkan suatu solusi harus benar-benar diperhatikan dalam pembangkitan setiap individunya (Suwirmayanti, 2014).

2. Nilai Fitness

Nilai *fitness* menyatakan nilai dari fungsi tujuan. Tujuan dari algoritma genetika sendiri yaitu memaksimalkan nilai *fitness*. Apabila yang dicari merupakan nilai maksimal, maka nilai *fitness* merupakan nilai dari fungsi itu sendiri. Tetapi jika yang dicari nilai minimal, maka nilai *fitness* merupakan *invers* dari nilai fungsi itu sendiri.

$$f = \frac{1}{(h + a)}$$

Dimana:

f : *fitness*

h : fungsi h

a : bilangan yang dianggap sangat kecil dan bervariasi sesuai dengan masalah yang akan diselesaikan (Suwirmayanti, 2014).

3. Seleksi

Proses seleksi merupakan proses untuk menentukan individu-individu mana saja yang akan dipilih untuk dilakukan *crossover*. Seleksi dilakukan untuk mendapatkan calon induk yang baik, induk yang baik akan menghasilkan keturunan yang baik. Semakin tinggi nilai *fitness* suatu individu maka akan semakin besar kemungkinannya untuk terpilih. Seleksi dapat dilakukan dengan menggunakan tiga macam metode, yaitu : Metode *roulette-wheel* (roda *roulette*), Metode seleksi rangking, Metode seleksi turnamen (Suwirmayanti, 2014).

4. Pindah Silang (*Crossover*)

Pindah silang atau *crossover* merupakan salah satu proses dalam algoritma genetika yang memiliki peran yang paling penting karena di dalamnya terdapat proses perkawinan (persilangan) gen antara dua individu (orang tua) secara acak yang menghasilkan dua individu baru (keturunan) pada generasi berikutnya. Beberapa teknik atau metode *crossover* yaitu Pindah silang satu titik (*one-point crossover*), Pindah silang dua titik (*two-point crossover*), Pindah silang banyak titik (*N-point crossover*) (Utami, dkk. 2014).

5. Mutasi

Mutasi merupakan suatu proses dalam algoritma genetika untuk menciptakan individu baru dengan melakukan modifikasi satu atau lebih gen dalam individu yang sama. Mutasi berfungsi untuk menggantikan gen yang hilang dari populasi selama proses seleksi serta menyediakan gen yang tidak ada dalam populasi awal. Sehingga mutasi akan meningkatkan variasi populasi. Metode mutasi yang biasanya digunakan yaitu: inversi, *Random*, *Shift*, *Swap* (Suwirmayanti, 2014).

Cost Function Criterion

Clustering criterion yang juga biasa disebut sebagai metode pengukuran hasil gerombol. *Clustering criterion* inilah yang akan dijadikan nilai *fitness* dari setiap kromosom yang telah dievaluasi. Semakin tinggi nilai *fitness*, maka akan semakin bagus kromosom yang dievaluasi tersebut. *Clustering criterion* yang digunakan adalah *cost function* atau biaya yang dihabiskan untuk menempatkan *object* pada gerombol yang bersesuaian. Adapun persamaan *cost function* untuk data campuran numerik dan kategorikal adalah sebagai berikut:

$$Cost = \sum_{i=1}^k (Cost_i^r + Cost_i^c) = \sum_{i=1}^k Cost_i^r + \sum_{i=1}^k Cost_i^c = Cost^r + Cost^c$$

Persamaan diatas adalah *cost function* untuk *clustering dataset* dengan atribut bernilai numerik dan kategorikal. Karena $Cost^r$ dan $Cost^c$ adalah non-negatif, minimalisasi *Cost* bisa dilakukan dengan meminimalkan $Cost^r$ dan $Cost^c$, *total cost* dari atribut numerik dan kategorikal untuk semua gerombol (Nooraeni, 2016).

3. METODOLOGI PENELITIAN

Lokasi Penelitian

Lokasi penelitian di kantor Badan Pusat Statistika (BPS) Provinsi Sulawesi Tengah dan tempat penelitian ini dilakukan di Laboratorium Statistika Terapan, Program Studi Statistika, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Tadulako.

Populasi dan Sampel Penelitian

Populasi pada penelitian ini adalah data sarana prasarana dan tenaga kesehatan di provinsi Sulawesi Tengah tahun 2018 dan sampel yang digunakan pada penelitian ini adalah data sarana prasarana dan tenaga kesehatan yang berada di setiap desa di Kabupaten Poso tahun 2018.

Prosedur Pengambilan Data

Data yang digunakan pada penelitian ini berupa data sekunder yang berasal dari Badan Pusat Statistika (BPS) provinsi Sulawesi Tengah. Variabel penelitian yang digunakan adalah jumlah tenaga kesehatan dan ketersediaan sarana prasarana di kabupaten Poso tahun 2018.

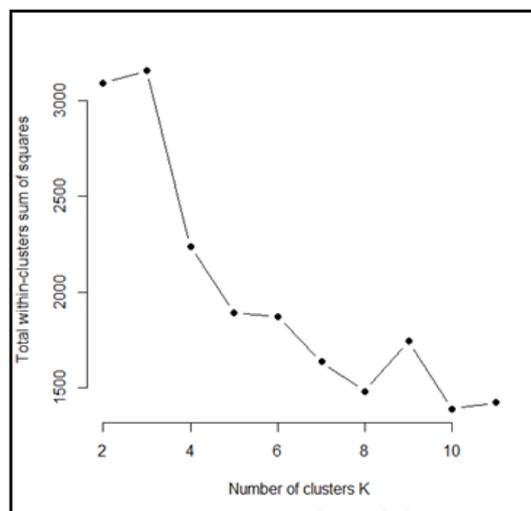
Analisis Data

1. Menentukan Jumlah K yang akan digunakan.
2. Menentukan inialisasi populasi awal, meliputi jumlah gen dalam kromosom dan jumlah kromosom dalam individu.
3. Melakukan proses algoritma k -prototype untuk setiap kromosom pada populasi.
4. Menghitung nilai $fitness$ dari tiap kromosom dalam populasi berdasarkan nilai $cost$ function.
5. Memilih kromosom berdasarkan nilai $fitness$.
6. Melakukan seleksi, $crossover$, dan mutasi untuk mendapatkan keturunan juga diperoleh populasi baru..
7. Kembali ke langkah 3 hingga kriteria pengoptimasian yang ditentukan terpenuhi.
8. Menginterpretasikan hasil gerombol.

4. HASIL DAN PEMBAHASAN

Penentuan Jumlah Gerombol (K)

Sebelum melakukan inialisasi populasi awal terlebih dahulu menentukan nilai K yang akan digunakan menggunakan metode $elbow$ dengan rentang nilai K yang diuji yaitu $K = 2$ sampai dengan $K = 11$. Dimana penentuan jumlah K dapat dilihat pada Gambar 1.



Gambar 1 Grafik Penentuan Jumlah K

Pada gambar 1 dari jumlah $cluster$ $K = 2$ hingga $K = 11$ dapat terlihat garis yang mengalami patahan membentuk $elbow$ atau siku pada saat $K = 5$, $K = 8$ dan $K = 10$. Berdasarkan efisiensi waktu pengolahan dan pertimbangan kecukupan jumlah $cluster$ maka nilai K yang ditetapkan adalah $K = 8$.

Inisialisasi Populasi Awal

Proses inialisasi populasi awal ini digunakan untuk menentukan sejumlah kromosom awal yang akan digunakan untuk proses selanjutnya. Pada penelitian ini menggunakan 12 variabel

dan 142 data desa di Kabupaten Poso, maka dalam membentuk kromosom individu panjang gen adalah sama dengan jumlah K dari proses penggerombolan. Dimana masing-masing nilai yang ada pada gen akan mewakili nomor data desa pada proses penggerombolan. Jadi nilai yang ada pada gen adalah nomor data desa yang terpilih secara acak dari nomor 1 sampai 142 sebanyak K gerombol. Jika jumlah kromosom yang akan dibentuk adalah 10 dan $K = 8$ maka akan ada $10 \times K$ yaitu 80 desa yang akan terpilih pada populasi awal ini.

Tabel 1 Inisialisasi Populasi Awal

Kromosom	Nilai Gen
1	[104, 120, 70, 54, 35, 16, 139, 78]
2	[31, 63, 142, 70, 49, 90, 51, 48]
3	[76, 105, 31, 58, 37, 87, 26, 117]
4	[107, 95, 87, 52, 74, 120, 80, 114]
5	[45, 100, 38, 83, 67, 37, 77, 124]
6	[129, 39, 46, 138, 86, 142, 64, 55]
7	[94, 22, 81, 34, 133, 83, 71, 55]
8	[125, 52, 41, 24, 139, 67, 35, 30]
9	[96, 7, 99, 49, 57, 113, 125, 39]
10	[137, 103, 97, 8, 55, 66, 77, 95]

Dari hasil Tabel 1 dapat diketahui bahwa pada proses inisialisasi populasi awal yang dilakukan terdapat 10 kromosom dengan nilai gen yang dihasilkan secara acak.

Penggerombolan Menggunakan *K*-prototype Generasi Pertama

Pada proses penggerombolan menggunakan *k*-prototype untuk setiap kromosom pada populasi bertujuan untuk menghasilkan nilai *cost function* dalam hal ini yang digunakan adalah total *within* dari gerombol. Pada generasi pertama terdapat 10 kromosom dengan nilai total *within* dari masing-masing kromosom disajikan pada Tabel 2.

Tabel 2 Nilai Total *Within* dari Kromosom Generasi Pertama

Kromosom	Total <i>Within</i>
1	1441,45
2	1524,377
3	1529,809
4	1632,912
5	1500,88
6	1671,359
7	1530,613
8	1716,927
9	1621,769
10	1487,933

Evaluasi Fitness Generasi Pertama

Setelah didapatkan nilai total *within* dari masing-masing kromosom, selanjutnya akan menghitung nilai *fitness* dari masing-masing kromosom menggunakan persamaan $f = \frac{1}{h+a}$.

Pada generasi pertama terdapat 10 kromosom dengan nilai *fitness* dari masing-masing kromosom disajikan pada Tabel 3.

Tabel 3 Nilai *Fitness* dari Kromosom Generasi Pertama

Kromosom	<i>Fitness</i>
1	0,0006937
2	0,000656
3	0,000654
4	0,000612
5	0,000666
6	0,000598
7	0,000653
8	0,000582
9	0,000617
10	0,000672

Seleksi Generasi Pertama

Proses seleksi bertujuan untuk mendapatkan calon kromosom induk yang baik untuk menjalani proses pindah silang (*crossover*) dan mutasi. Metode yang akan digunakan dalam proses seleksi adalah *roulette-wheel*. Pada generasi pertama terjadi perubahan kromosom setelah melakukan seleksi *roulette-wheel* dimana perubahan dari masing-masing kromosom disajikan pada Tabel 4 dan hasil seleksi dari kromosom 1 sampai 10 disajikan pada Tabel 5. Setelah melakukan seleksi maka kromosom induk yang terpilih adalah kromosom dengan nilai *fitness* terbesar yaitu kromosom (1) dan kromosom (10) dengan nilai *fitness* yang dihasilkan sebesar 0,0006937 dan 0,000672.

Tabel 4 Kromosom Sebelum dan Sesudah Seleksi Generasi Pertama

Kromosom Sebelum Seleksi	Kromosom setelah Seleksi
Kromosom 1	Kromosom 6
Kromosom 2	Kromosom 5
Kromosom 3	Kromosom 10
Kromosom 4	Kromosom 8
Kromosom 5	Kromosom 1
Kromosom 6	Kromosom 6
Kromosom 7	Kromosom 7
Kromosom 8	Kromosom 9
Kromosom 9	Kromosom 3
Kromosom 10	Kromosom 1

Tabel 5 Hasil Seleksi Kromosom Generasi Pertama

Kromosom	Nilai Gen setelah Seleksi
1	[129, 39, 46, 138, 86, 142, 64, 55]
2	[45, 100, 38, 83, 67, 37, 77, 124]
3	[137, 103, 97, 8, 55, 66, 77, 95]
4	[125, 52, 41, 24, 139, 67, 35, 30]
5	[104, 120, 70, 54, 35, 16, 139, 78]
6	[129, 39, 46, 138, 86, 142, 64, 55]
7	[94, 22, 81, 34, 133, 83, 71, 55]
8	[96, 7, 99, 49, 57, 113, 125, 39]
9	[76, 105, 31, 58, 37, 87, 26, 117]
10	[104, 120, 70, 54, 35, 16, 139, 78]

Pindah Silang (Crossover) Generasi Pertama

Setelah melakukan proses seleksi untuk mendapatkan kromosom induk, maka selanjutnya dilakukan proses pindah silang (*crossover*) untuk kromosom induk agar menghasilkan keturunan baru dengan melakukan pindah silang satu titik (*one-point crossover*) dengan posisi titik potong pindah silang (*crossover*) dipilih menggunakan bilangan acak. Kromosom induk yang akan melakukan proses pindah silang (*crossover*) adalah kromosom (1) dan kromosom (10).

Posisi titik potong pindah silang dipilih menggunakan bilangan acak 1-7, dimana titik potong untuk kromosom (1) yaitu 4 dan titik potong untuk kromosom (10) yaitu 6. Hasil dari pindah silang disajikan pada Tabel 6.

Tabel 6 Hasil Pindah Silang (Crossover) Generasi Pertama

Offspring	Sebelum Pindah Silang	Hasil Pindah Silang
Kromosom 1	129, 39, 46, 138, 86, 142, 64, 55	[129, 39, 46, 138, 35, 16, 139, 78]
Kromosom 10	104, 120, 70, 54, 35, 16, 139, 78	[104, 120, 70, 54, 35, 16, 64, 55]

Mutasi Generasi Pertama

Proses mutasi merupakan proses yang dilakukan dengan cara mengganti satu gen yang terpilih secara acak dengan satu nilai gen baru yang didapat secara acak. Adapun total gen dalam proses mutasi ini sebanyak 80 gen yang akan terpilih secara acak untuk digantikan posisinya dengan nilai gen baru yang dibangkitkan secara acak antara 1 sampai dengan 142 dan jumlah mutasi sebanyak 8. Hasil mutasi kromosom pada generasi pertama disajikan pada Tabel 7.

Tabel 7 Hasil Mutasi Kromosom Generasi Pertama

Kromosom	Nilai Gen Sebelum Mutasi	Nilai Gen Setelah Mutasi
1	[129,39,46,138,35,16,139,78]	[129,39,46,138,35,16,139,78]
2	[45,100,38,83,67, 37 ,77,124]	[45,100,38,83,67, 117 ,77,124]
3	[137, 103 ,97,8,55,66,77,95]	[137, 110 ,97,8,55,66,77,95]
4	[125,52,41,24,139,67, 35 ,30]	[125,52,41,24,139,67, 32 ,30]
5	[104, 120 ,70,54,35,16,139,78]	[104, 62 ,70,54,35,16,139,78]
6	[129,39,46,138, 86 ,142,64,55]	[129,39,46,138, 18 ,142,64,55]
7	[94,22, 81 ,34,133,83,71,55]	[94,22, 135 ,34,133,83,71,55]
8	[96 ,7,99,49,57,113,125,39]	[36 ,7,99,49,57,113,125,39]
9	[76,105,31,58,37, 87 ,26,117]	[76,105,31,58,37, 52 ,26,117]
10	[104,120,70,54,35,16,64,55]	[104,120,70,54,35,16,64,55]

Dari hasil Tabel 7 terdapat nilai gen dari masing-masing kromosom sebelum dan sesudah mutasi. Posisi gen yang mengalami mutasi dibangkitkan secara acak. Nilai gen pada posisi yang terpilih juga mengalami perubahan dimana nilai gen yang digantikan berasal dari pembangkitan secara acak antara 1-142, dimana nilai gen yang dibangkitkan adalah 117, 110, 32, 62, 18, 135, 36 dan 52. Proses ini tentu dapat mengubah nilai *fitness* kromosom-kromosom tersebut menjadi lebih rendah maupun menjadi lebih tinggi. Setelah dilakukan proses mutasi maka telah diselesaikan satu generasi dan kromosom-kromosom dari hasil mutasi di atas akan dijadikan individu baru yang akan di proses kembali pada generasi selanjutnya.

Proses penggerombolan dengan algoritma genetika dilakukan terus menerus sampai dengan generasi ke sepuluh dengan mengulangi tiap tahapan seperti pada proses generasi pertama. Pada generasi kesepuluh terdapat 10 kromosom dengan nilai *fitness* terbesar adalah 0,000701 terdapat pada kromosom (1) yang merupakan kromosom terbaik dengan nilai gen pada kromosom (1) adalah 36, 7, 99, 49, 69, 104, 105, 110 yang akan digunakan mejadi inisial *center* pada proses gerombol dengan algoritma *k-prototype*.

Hasil Penggerombolan Menggunakan *K-prototype*

Algoritma *k-prototype* merupakan suatu metode gerombol yang dapat melakukan penggerombolan pada data dengan tipe campuran (numerik dan kategorik). Untuk tabel hasil penggerombolan disajikan sebagai berikut.

Tabel 8 Hasil Penggerombolan *K-prototype*

Gerombol Ke-i	Jumlah Anggota (Desa)	Nama Desa
Gerombol 1	3	Pendolo, Pasir Putih, Wuasa
Gerombol 2	16	Panda Jaya, Salukaia, Toinasa, Tiu, Poleganyara, Sangira, Alitupu, Bariri, Tamadue, Maholo, Bega, Masamba, Tangkura, Padalembara, Silanca, Malei
Gerombol 3	15	Bangun Jaya, Salindu, Tolambo, Tindoli, Kageroa, Uelincu, Kuku, Sedoa, Rompo, Tiwaa, Lape, Malitu, Patiwunga, Betalemba, Membuke
Gerombol 4	6	Gintu, Petiro, Sulewana, Doda, Tambarana, Tagolu
Gerombol 5	32	Maya Jaya, Taipa, Uranosari, Runde, Buyumpondoli, Peura, Leboni, Pancasila I, Kamba, Matialemba, Magapu, Saojo, Lena, Panjoka, Kaduwaa, Bumi Banyusari, Hanggira, Lempe, Torire, Baliura, Katu, Siliwanga, Toini, Masani, Towu, Ueralulu, Kilo, Tumora, Maranda, Tampemadoro, Toyado, Labuan
Gerombol 6	17	Panjo, Meko, Singkona, Korobono, Bulili, Bewa, Tonusu, Dulumai, Olumokunde, Sangginora, Dewua, Tri Mulya, Kawende, Kalora, Tobe, Watuawu, Tongko
Gerombol 7	16	Mayoa, Uelene, Pandayora, Boe, Betue, Watutau, Betania, Saatu, Tokorondo, Lanto Jaya, Pantangolemba, Taunca, Bakti Agung, Tambaro, Sepe, Labuadago
Gerombol 8	37	Maya Sari, Bancea, Owini, Barati, Tokilo, Amporiwo, Wayura, Badangkaia, Bakekau, Pada, Bomba, Tuare, Tomehipi, Lengkeka, Kolori, Lelio, Mayakeli, Soe, Wera, Masewe, Taripa, Ompo Didiri, Kelei, Kancuu, Dodolo, Watumaeta, Winowanga, Mekar Sari, Kalemago, Talabosa, Wanga, Pinedapa, Pandiri, Sintuwulemba, Maliwuko, Bategencu, Ratoombu

Karakteristik setiap gerombol dari masing-masing variabel kategorik dan numerik sebagai berikut :

Tabel 9 Kararteristik Setiap Gerombol Untuk Variabel Kategorik

Gerombol	Variabel Kategorik							
	X1	X2	X3	X4	X5	X6	X7	X8
1	1	2	1	2	1	2	1	2
2	2	2	2	2	2	1	2	2
3	2	2	2	2	2	1	1	2
4	1	2	2	2	2	2	2	2
5	2	2	2	2	2	2	1	2
6	2	2	2	2	1	1	2	2
7	2	2	2	2	2	2	1	2
8	2	2	2	2	2	1	2	2

Dari Tabel 9 dapat diketahui bahwa gerombol 1 untuk variabel kategorik memiliki karakteristik desa dimana tersedia puskesmas dengan rawat inap (X1), tempat praktek dokter (X3), tempat praktek bidan (X5) dan polindes (X7), sedangkan untuk puskesmas tanpa rawat inap (X2), rumah bersalin (X4), poskesdes (X6) dan apotek (X8) tidak tersedia. Gerombol 2 untuk variabel kategorik memiliki karakteristik desa dimana tersedia poskesdes (X6), sedangkan untuk puskesmas dengan rawat inap (X1), puskesmas tanpa rawat inap (X2), tempat praktek dokter (X3), rumah bersalin (X4), tempat praktek bidan (X5), polindes (X7) dan apotek (X8) tidak tersedia. Gerombol 3 untuk variabel kategorik memiliki karakteristik desa dimana tersedia poskesdes (X6) dan polindes (X7) sedangkan untuk puskesmas dengan rawat inap (X1), puskesmas tanpa rawat inap (X2), tempat praktek dokter (X3), rumah bersalin (X4), tempat praktek bidan (X5) dan apotek (X8) tidak tersedia. Gerombol 4 untuk variabel kategorik memiliki karakteristik desa dimana tersedia puskesmas dengan rawat inap (X1) sedangkan puskesmas tanpa rawat inap (X2), tempat praktek dokter (X3), rumah bersalin (X4), tempat praktek bidan (X5), poskesdes (X6), polindes (X7) dan apotek (X8) tidak tersedia. Gerombol 5 untuk variabel kategorik memiliki karakteristik desa dimana tersedia polindes (X7) sedangkan puskesmas dengan rawat inap (X1), puskesmas tanpa rawat inap (X2), tempat praktek dokter (X3), rumah bersalin (X4), tempat praktek bidan (X5), poskesdes (X6) dan apotek (X8) tidak tersedia. Gerombol 6 untuk variabel kategorik memiliki karakteristik desa dimana tersedia tempat praktek bidan (X5) dan poskesdes (X6) sedangkan puskesmas dengan rawat inap (X1), puskesmas tanpa rawat inap (X2), tempat praktek dokter (X3), rumah bersalin (X4), polindes (X7) dan apotek (X8) tidak tersedia. Gerombol 7 untuk variabel kategorik memiliki karakteristik desa dimana tersedia polindes (X7) sedangkan puskesmas dengan rawat inap (X1), puskesmas tanpa rawat inap (X2), tempat praktek dokter (X3), rumah bersalin (X4), tempat praktek bidan (X5), poskesdes (X6) dan apotek (X8) tidak tersedia. Gerombol 8 untuk variabel kategorik memiliki karakteristik desa dimana tersedia poskesdes (X6) sedangkan puskesmas dengan rawat inap (X1), puskesmas tanpa rawat inap (X2), tempat praktek dokter (X3), rumah bersalin (X4), tempat praktek bidan (X5), polindes (X7) dan apotek (X8) tidak tersedia. Dan untuk Kararteristik Setiap Gerombol Untuk Variabel Numerik disajikan pada Tabel 10.

Tabel 10 Kararteristik Setiap Gerombol Untuk Variabel Numerik

Gerombol	Variabel Numerik			
	X9	X10	X11	X12
1	1	1	3	16
2	0	0	2	6
3	0	0	1	1
4	1	0	4	5
5	0	0	2	1
6	0	0	2	3
7	0	0	2	6
8	0	0	1	2

Dari Tabel 10 dapat diketahui bahwa gerombol 1 untuk variabel numerik memiliki karakteristik desa dimana rata-rata tenaga dokter sebesar 1, rata-rata tenaga dokter gigi sebesar 1, rata-rata tenaga bidan sebesar 3 dan rata-rata tenaga kesehatan lain sebesar 16. Gerombol 2 untuk variabel numerik memiliki karakteristik desa dimana rata-rata tenaga dokter sebesar 0, rata-rata tenaga dokter gigi sebesar 0, rata-rata tenaga bidan sebesar 2 dan rata-rata tenaga kesehatan lain sebesar 6. Gerombol 3 untuk variabel numerik memiliki karakteristik desa dimana rata-rata tenaga dokter sebesar 0, rata-rata tenaga dokter gigi sebesar 0, rata-rata tenaga bidan sebesar 1 dan rata-rata tenaga kesehatan lain sebesar 1. Gerombol 4 untuk variabel numerik memiliki karakteristik desa dimana rata-rata tenaga dokter sebesar 1, rata-rata tenaga dokter gigi sebesar 0, rata-rata tenaga bidan sebesar 4, dan rata-rata tenaga kesehatan lain sebesar 5. Gerombol 5 untuk variabel numerik memiliki karakteristik desa dimana rata-rata tenaga dokter sebesar 0, rata-rata tenaga dokter gigi sebesar 0, rata-rata tenaga bidan sebesar 2 dan rata-rata tenaga kesehatan lain sebesar 1. Gerombol 6 untuk variabel numerik memiliki karakteristik desa dimana rata-rata tenaga dokter sebesar 0, rata-rata tenaga dokter gigi sebesar 0, rata-rata tenaga bidan sebesar 2 dan rata-rata tenaga kesehatan lain sebesar 3. Gerombol 7 untuk variabel numerik memiliki karakteristik desa dimana rata-rata tenaga dokter sebesar 0, rata-rata tenaga dokter gigi sebesar 0, rata-rata tenaga bidan sebesar 2 dan rata-rata tenaga kesehatan lain sebesar 6. Gerombol 8 untuk variabel numerik memiliki karakteristik desa dimana rata-rata tenaga dokter sebesar 0, rata-rata tenaga dokter gigi sebesar 0, rata-rata tenaga bidan sebesar 1 dan rata-rata tenaga kesehatan lain sebesar 2.

5. KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil dan pembahasan yang telah diuraikan sebelumnya, maka dapat diambil kesimpulan bahwa hasil pegerombol desa di Kabupaten Poso berdasarkan sarana prasarana dan tenaga kesehatan menggunakan metode *k-prototype* yang dioptimalkan dengan algoritma genetika diperoleh 8 gerombol dengan masing-masing jumlah desa pada tiap gerombol yaitu 3 desa pada gerombol pertama, 16 desa pada gerombol kedua, 15 desa pada gerombol ketiga, 6 desa pada gerombol keempat, 32 desa pada gerombol kelima, 17 desa pada gerombol keenam, 16 desa pada gerombol ketujuh dan 37 desa pada gerombol kedelapan.

Saran

Dari hasil penelitian ini, penulis menyarankan untuk penelitian selanjutnya sebaiknya melakukan proses generasi lebih banyak atau sampai mencapai konvergen agar pusat gerombol yang dihasilkan lebih baik lagi dan juga disarankan untuk dapat menggunakan metode gerombol yang berbeda dan metode optimasi yang berbeda agar dapat menjadi pembandingan sehingga memungkinkan mendapatkan hasil yang lebih baik.

DAFTAR PUSTAKA

- Ahmad, A dan Lipika, D. (2007). A K-Mean Clustering Algorithm For Mixed Numeric And Categorical Data, *Data & Knowledge Engineering*, vol. 63, PP. 503-527.
- [BPS] Badan Pusat Statistik. (2019). *Statistik Daerah Kabupaten Poso 2019*. Kabupaten Poso: BPS.
- Desiani, A dan Arhami, M. (2006). *Konsep Kecerdasan Buatan*. Yogyakarta: Andi Offset.
- [Dinkes] Dinas Kesehatan. (2018). *Profil Kesehatan Provinsi Sulawesi Tengah Tahun 2018*. Palu. Dinkes Sulteng.
- Han, J dan Kamber. (2006). *Data Mining Concepts and Techniques*, 2nd ed. San Fransisco, United States of America: Dianne Cerra.
- Huang, Z. (1997). *Clustering Large Data Sets with Mixed Numeric and Categorical Values*. 1997.
- Huang, Z. (1998). *Extensions To The K-Means Algorithm for Clustering Large Data Sets with Categorical Values*, *Data Mining and Knowledge Discovery* 2, 283-304.
- Madhultha, T.S. (2012). *An Overview On Clustering Methods*. *IOSR Journal Of Engineering*, II(4), PP.719-725.
- Nooraeni, R. (2016). *Kombinasi Algoritma Cluster K-Prototype dan Algoritma Genetika untuk data Bertipe Campuran*. Jakarta: Sekolah Tinggi Ilmu Statistik.
- Suwirmayanti, P. (2014). *Optimasi Pusat Cluster K-Prototype dengan Algoritma Genetika*. Denpasar: Universitas Udayana.
- Utami, P.Y. Suherly, C dan Ilhamsyah. (2014). *Aplikasi Pencarian Rute Terpendek Menggunakan Algoritma Genetika*. *Jurnal Coding Sistem Komputer Universitas Tanjungpura*.

LAMPIRAN

Lampiran 1. Sintaks dan *Output* Penentuan Jumlah K

```

library(openxlsx)
library(clustMixType)
data=read.xlsx("E:/DATA KU FIX.xlsx")

data$X1<-as.factor(data$X1)
data$X2<-as.factor(data$X2)
data$X3<-as.factor(data$X3)
data$X4<-as.factor(data$X4)
data$X5<-as.factor(data$X5)
data$X6<-as.factor(data$X6)
data$X7<-as.factor(data$X7)
data$X8<-as.factor(data$X8)
data$X9<-as.numeric(data$X9)
data$X10<-as.numeric(data$X10)
data$X11<-as.numeric(data$X11)
data$X12<-as.numeric(data$X12)

fac=cbind(data$X1,data$X2,data$X3,data$X4,data$X5,data$X6,data$X7,data$X8)
num=cbind(data$X9,data$X10,data$X11,data$X12)

x<-data.frame(data$X1,data$X2,data$X3,data$X4,data$X5,data$X6,data$X7,data$X8,data$X9,data$X10,data$X11,data$X12)

set.seed(123)
wss<-sapply(2:11,
function(k){kproto(x,k)$tot.withinss})
wss
plot(2:11,wss,type="b",pch=19, frame=FALSE,
xlab="Number of clusters K",
ylab="Total within-clusters sum of squares")

```

Lampiran 2. Sintaks dan *Output* Inisialisasi Populasi Awal

```

library(ChemometricsWithR)
GA.init.pop(10,142,8,8)

```

Lampiran 3. Sintaks Nilai Total *Within* dari Kromosom dan Penggerombolan Menggunakan *K-prototype*

```

k110<-read.delim("clipboard")

set.seed(123)
k110$X1<-as.factor(k110$X1)
k110$X2<-as.factor(k110$X2)
k110$X3<-as.factor(k110$X3)
k110$X4<-as.factor(k110$X4)
k110$X5<-as.factor(k110$X5)
k110$X6<-as.factor(k110$X6)
k110$X7<-as.factor(k110$X7)
k110$X8<-as.factor(k110$X8)
k110$X9<-as.numeric(k110$X9)
k110$X10<-as.numeric(k110$X10)
k110$X11<-as.numeric(k110$X11)
k110$X12<-as.numeric(k110$X12)

pro110<-kproto(x,8,centers=k110)
ds110<-pro110$cluster
cf110<-pro110$tot.withinss

```