

KLASTERISASI PENERIMA DANA BANTUAN PROGRAM KELUARGA HARAPAN MENGGUNAKAN METODE K-MEANS PADA DESA GEREBA

Ninda Siti Paridah¹, Martanto²

¹ Program Studi Komputerisasi Akuntansi, STMIK IKMI Cirebon

² Program Studi Manajemen Informatika, STMIK IKMI Cirebon
nindasitifaridah69@gmail.com

ABSTRAK

Salah satu unsur yang harus diperhatikan pemerintah untuk mencapai kesejahteraan penduduk adalah persoalan kemiskinan. Salah satu upaya penanggulangan kemiskinan, pemerintah memberikan bantuan kepada masyarakat yang diberi nama program keluarga harapan (PKH). Desa yang menerima bantuan dana ini adalah Desa Gereba. Dari informasi yang diperoleh dari penerima PKH, terdapat tantangan dalam memproses data untuk menentukan masyarakat miskin yang harus menjadi prioritas dalam menerima bantuan, juga kesulitan dalam mendistribusikan bantuan secara merata di berbagai dusun di Desa Gereba dikarenakan banyak masyarakat merasa dirinya berhak menerima bantuan PKH, walaupun tidak memenuhi persyaratan yang ada. Tujuan penulisan tugas akhir ini untuk membantu pemerintahan desa dalam mengelompokkan data penduduk yang layak menerima bantuan menggunakan algoritma *K-means clustering*. *K-Means* merupakan sebuah metode dalam data mining yang memiliki kemampuan untuk melakukan pengelompokan atau *Clustering* data, dimana entitas yang memiliki atribut serupa dikelompokkan ke dalam satu Hasil penelitian didapatkan yaitu menggunakan algoritma *K-Means* dibagi menjadi 2 cluster pada cluster 0 berjumlah 163 items, cluster 1 berjumlah 167 items. Cluster optimum dan nilai *DBI* terdapat pada cluster 2 tanpa menggunakan operator *Normalize* dengan nilai *Davies Bouldin Index* atau *DBI* yaitu 0.500. Maka dusun yang menjadi prioritas yaitu Dusun Ciawitali.

Kata Kunci : *DBI, Clustering, K-Means, Komponen, PKH*

1. PENDAHULUAN

Kemajuan suatu daerah bisa dinilai dari tingkat kesejahteraan penduduknya. Faktor penting dalam mencapai kesejahteraan yaitu penanggulangan kemiskinan, yang menghambat proses pembangunan, terlihat dari keterbatasan, ketidakmampuan, dan kekurangan yang banyak dialami. Keterbatasan mencakup kesulitan dalam mencapai kebebasan hidup sesuai dengan harapan hidup, kesulitan dalam akses pendidikan, akses yang terbatas terhadap layanan kesehatan yang memadai, serta kekurangan dalam pemenuhan kebutuhan dasar seperti sandang dan pangan. [1].

Untuk mengatasi isu kemiskinan pemerintah telah melakukan berbagai upaya dengan mengadakan bantuan kepada masyarakat diantaranya dengan meluncurkan program khusus yang diberi nama program keluarga harapan (PKH). PKH merupakan sebuah program bantuan sosial dengan syarat yang diberikan kepada keluarga atau individu yang tergolong dalam kategori miskin dan rentan, yang terdaftar dalam basis data terpadu untuk penanganan fakir miskin. Data ini diolah oleh pusat data dan informasi kesejahteraan sosial dan digunakan untuk menentukan keluarga yang menjadi penerima manfaat dari program PKH. [2].

Bantuan yang disalurkan melalui PKH memiliki persyaratan yang harus dipenuhi terutama untuk rumah tangga miskin dan memiliki komponen antara lain yaitu: anak sekolah, balita, ibu hamil, penyandang disabilitas, dan lansia, rumah tangga penerima bantuan PKH secara nominal sesuai dengan kriteria komponen.

Bantuan ini memiliki syarat tertentu, seperti kehadiran minimal 80% untuk anak sekolah, kunjungan bulanan ke puskesmas untuk balita, ibu hamil, dan individu dengan disabilitas untuk pemeriksaan kesehatan dan imunisasi. Lansia diwajibkan ke posyandu (pos pelayanan terpadu) dan melakukan pemeriksaan kesehatan setiap bulannya.

Salah satu desa yang mendapatkan alokasi dana bantuan PKH yaitu Desa Gereba. Desa Gereba merupakan salah satu desa di Kabupaten Ciamis yang terdiri dari 4 dusun yaitu Dusun Ciawitali, Dusun Gereba, Dusun Nanggerang, Dusun Ciroyom dengan jumlah KK 1.257, dari jumlah KK tersebut penerima bantuan PKH di Desa Gereba yaitu sebanyak 330 KK. Untuk mengoptimalkan bantuan dana tersebut tentunya harus disesuaikan dengan prasyarat penerima PKH, namun penerimaan bantuan dana PKH dinilai oleh beberapa masyarakat tidak tepat sasaran sehingga terjadi kecemburuan sosial antar masyarakat di Desa Gereba. *clustering*) data penduduk yang memenuhi syarat untuk menerima bantuan dengan akurat menggunakan algoritma *K-means clustering*.

Hasil pengklusteran akan disesuaikan dengan menggunakan *clustering* komponen penerima bantuan PKH sesuai prasyarat yang ditentukan, sehingga bantuan dapat diterima secara merata dan tepat sasaran kepada masyarakat miskin di Desa Gereba dapat terlaksana dengan baik, serta staff desa dapat mengetahui dusun mana yang memiliki jumlah masyarakat miskin terbanyak setiap dusunnya sehingga dapat mendapatkan perhatian lebih untuk tahun-tahun yang akan datang.

Dari latar belakang masalah yang ada maka diusulkan tugas akhir dengan judul “Klasterisasi Penerima Dana Bantuan Program Keluarga Harapan Menggunakan Metode K-Means Pada Desa Gereba”.

2. TINJAUAN PUSTAKA

2.1. Program Keluarga Harapan (PKH)

Menurut Peraturan Menteri No 1 Tahun 2018 Program Keluarga Harapan (PKH) adalah 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. Tujuan PKH menurut PERMENSOS No 1 Tahun 2018 yaitu:

1. Untuk meningkatkan kualitas hidup keluarga penerima manfaat melalui akses yang lebih baik ke layanan pendidikan, kesehatan, dan kesejahteraan sosial.
2. Mengurangi beban pengeluaran dan meningkatkan pendapatan keluarga yang miskin dan rentan.
3. Mendorong perubahan perilaku dan kemandirian keluarga penerima manfaat dalam mengakses layanan kesehatan, pendidikan, serta kesejahteraan sosial.
4. Mereduksi tingkat kemiskinan dan kesenjangan.

2.2. Data Mining

Menurut Alkhairi & Windarto dalam [3] data Mining merupakan serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui pengolahan data secara manual dari sebuah basis data. Pada tahun 1990-an, Data Mining mulai diperkenalkan sebagai metode yang tepat untuk mengekstrak pola dan informasi dari data guna menemukan keterkaitan antar data, memungkinkan pengelompokan dalam satu atau lebih cluster sehingga objek dalam satu cluster memiliki kesamaan yang signifikan satu sama lain. Data Mining merupakan bagian dari proses penemuan pengetahuan dari Knowledge Discovery in Databases, suatu proses pemahaman informasi dari basis data.

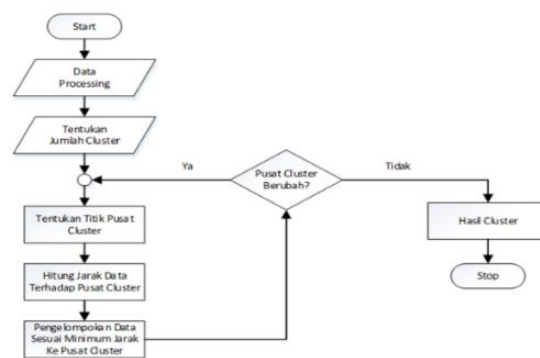
2.3. Klasterisasi (Clustering)

Klasterisasi atau Clustering termasuk kedalam salah satu teknik *unsupervised learning* dimana di dalamnya tidak perlu dilakukan pelatihan model atau bisa juga dikatakan dengan tidak diperlukan tahap pelatihan model, atau disebut juga tanpa fase pembelajaran. Clustering adalah proses yang membagi data menjadi beberapa kelompok atau cluster berdasarkan kesamaan atau ciri-ciri yang dimiliki oleh data tersebut. Data-data yang memiliki kemiripan akan berkumpul atau dikelompokkan di satu cluster yang sama. Sedangkan data-data yang memiliki karakteristik yang berbeda akan berkumpul atau dikelompokkan pada cluster yang berbeda. Tujuan dari

klasterisasi ini adalah pengelompokan data atau objek kedalam sebuah kelompok sehingga setiap kelompok akan berisi data-data yang memiliki karakteristik yang sangat mirip [4]

2.4. Metode K-Means

K-Means merupakan sebuah metode dalam data mining yang memiliki kemampuan untuk melakukan pengelompokan atau Clustering data, dimana entitas yang memiliki atribut serupa dikelompokkan ke dalam satu kelompok tunggal, sementara yang memiliki atribut yang berbeda ditempatkan dalam kelompok yang berbeda. [5]



Gambar 1. Langkah clustering menggunakan metode K-means

2.5. RapidMiner

RapidMiner adalah platform perangkat lunak yang kuat untuk ilmu data dan pembelajaran mesin. RapidMiner menyajikan berbagai perangkat untuk menyusun data, membuat model, mengevaluasi, dan menerapkan hasilnya. Dirancang untuk kemudahan penggunaan, platform ini memfasilitasi pembuatan dan pengujian model tanpa perlu keahlian pemrograman. RapidMiner menawarkan antarmuka *drag-and-drop* yang memungkinkan pengguna untuk membangun alur kerja untuk memproses dan menganalisis data. Ini mendukung beragam sumber data, termasuk file datar, basis data, dan platform big data seperti Hadoop dan Spark. Perangkat lunak ini juga mencakup beragam operator yang sudah dibangun, yang merupakan blok bangunan dari alur kerja, yang mencakup semua tahap proses data Mining, seperti pembersihan data, pemilihan fitur, dan pemodelan.[6]

3. METODE PENELITIAN

3.1. Sumber Data

Sumber data dalam penelitian ini diperoleh langsung dari salah satu staff Desa Gereba yaitu Bapak Usman Ali selaku penanggungjawab penyaluran PKH Desa Gereba, penelitian dilakukan di Kantor Desa Gereba pada tanggal 17 Oktober 2023.

3.1.1. Data Penelitian

Data yang digunakan dalam penelitian ini terdiri dari informasi yang bersumber langsung

(primer) serta informasi yang diperoleh dari sumber lain atau sebelumnya (sekunder). Data primer diperoleh dari hasil wawancara secara langsung dengan salah satu staff Desa Gereba yang juga merupakan penanggung jawab penyaluran dana bantuan PKH Desa Gereba. Data sekunder diperoleh dari informasi terkait Keluarga Penerima Manfaat (KPM) PKH yang berlokasi di area Desa Gereba. Data ini berisi nama dan alamat KPM PKH, komponen PKH serta jumlah bantuan yang diberikan sepanjang tahun 2023.

3.1.2. Metode Pengumpulan Data

Metode pengumpulan informasi yang digunakan dalam penelitian ini adalah:[7]

1. Observasi

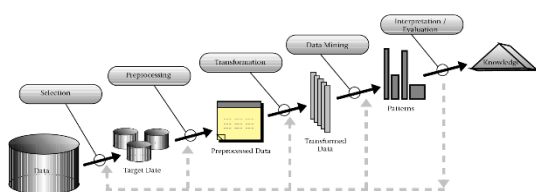
Dalam penelitian ini peneliti melakukan pengamatan langsung ke kantor Desa Gereba untuk mengetahui informasi penerima PKH serta permasalahan yang ada pada saat penyaluran bantuan PKH.

2. Wawancara

Dalam penelitian ini wawancara dilakukan dengan salah seorang staff Desa Gereba yaitu Bapak Usman Ali selaku penanggung jawab penyalur bantuan PKH. Wawancara yang diterapkan menggunakan pertanyaan yang terbuka, bertujuan untuk mendapatkan informasi sebanyak mungkin tentang penerima bantuan PKH dan masalah yang muncul dalam penyalurannya. Tujuannya adalah mendapatkan data yang detail tanpa kehilangan fokus pada topik utama serta tujuan penelitian yang sebenarnya.

3.2. Tahapan Perancangan

Tahapan-tahapan yang dilakukan dalam penelitian adalah menggunakan perancangan KDD. Menurut Tomar & Agarwal dalam [8] Data Mining adalah bagian integral dari enam proses dalam Knowledge Discovery in Databases (KDD). KDD merupakan serangkaian langkah untuk mengeksplorasi informasi yang lebih berharga, lebih mudah dipahami, dan baru dari data yang besar dan kompleks.[9] Proses KDD melibatkan interpretasi hasil dari kumpulan data dengan mengintegrasikan berbagai ilmu. Dimulai dengan menetapkan tujuan, proses KDD berakhir dengan evaluasi. [10]



Gambar 2. Tahapan KDD

Dari Gambar 2. dideskripsikan tahapan-tahapan dari KDD.

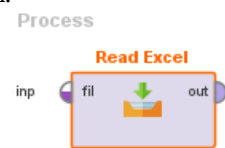
4. HASIL DAN PEMBAHASAN

4.1. Mengimplementasikan Algoritma K-Means

Dalam pengamatan ini menggunakan tahapan KDD pada pengelompokan data penerima bantuan PKH, adapun tahapan alur KDD yaitu :

4.1.1. Data Selection

Pada tahapan ini ada tujuh atribut yang didapat dari data penerima bantuan PKH Desa Gereba yaitu No, Nama Penerima, Kategori, Nama Pemilik, Nominal, dan Alamat dengan jumlah data sebanyak 330 record. Pada proses data selection tahapan pertama yang dilakukan yaitu mengimport seluruh data penerima bantuan PKH yang berformat excel (xlsx) pada aplikasi RapidMiner versi 9.9, apabila tombol pada operator read excel tidak lagi tanda seru warna kuning artinya operator telah berisi data dan siap untuk diolah.



Gambar 3. Operator read excel

Pada gambar 3 disajikan gambar operator read excel yang berfungsi untuk membaca data dengan format xlsx.

NO	NAMA PEN...	KATEGORI	NAMA PEM...	NOMINAL	Alamat
1	FALZAH HANAF	SMK	AKA DATI AH	375000	DUSUN CIAWITA
2	JERRE FOURADE	SMK	LILILAND	375000	DUSUN CIAWITA
3	ADAH	LANJISA	ADAH	800000	DUSUN CIAWITA
4	KEBIAH KOCERLURE	SMK	ACEH BERAWATI	800000	DUSUN CIAWITA
5	AFIRISA	SMK	ACEH BERAWATI	375000	DUSUN CIAWITA
6	AFIRISA	LANJISA	AFIRISA	800000	DUSUN CIAWITA
7	KORONG KOKALIA	BRU HAMIL	KORONG KOKALIA	750000	DUSUN CIAWITA
8	MAKESAPCHI	LANJISA	MAKESAPCHI	800000	DUSUN CIAWITA
9	CHULIA KOKAYAH	BRU HAMIL	CHULIA KOKAYAH	750000	DUSUN CIAWITA
10	LIA APRIELANI	SD	IMAS	225000	DUSUN CIAWITA
11	RIHANAH ADEHFI	SLITA	IMAS	750000	DUSUN CIAWITA
12	RECHU ALAMASTAH	SMK	FAIRFUD	800000	DUSUN CIAWITA

Gambar 4. Data excel pada operator read excel

Hasil data excel yang telah diimport menggunakan operator read excel disajikan pada gambar 4.



Gambar 5. Operator set role

Gambar 5. merupakan bentuk dari operator Set role yang digunakan untuk menetapkan field atau atribut yang akan dipergunakan sebagai id.

NO	NAMA PEN...	KATEGORI	NAMA PEM...	NOMINAL	Alamat
1	FALZAH HANAF	SMK	AKA DATI AH	375000	DUSUN CIAWITA
2	JERRE FOURADE	SMK	LILILAND	375000	DUSUN CIAWITA
3	ADAH	LANJISA	ADAH	800000	DUSUN CIAWITA
4	KEBIAH KOCERLURE	SMK	ACEH BERAWATI	800000	DUSUN CIAWITA
5	AFIRISA	SMK	ACEH BERAWATI	375000	DUSUN CIAWITA
6	AFIRISA	LANJISA	AFIRISA	800000	DUSUN CIAWITA
7	KORONG KOKALIA	BRU HAMIL	KORONG KOKALIA	750000	DUSUN CIAWITA
8	MAKESAPCHI	LANJISA	MAKESAPCHI	800000	DUSUN CIAWITA
9	CHULIA KOKAYAH	BRU HAMIL	CHULIA KOKAYAH	750000	DUSUN CIAWITA
10	LIA APRIELANI	SD	IMAS	225000	DUSUN CIAWITA
11	RIHANAH ADEHFI	SLITA	IMAS	750000	DUSUN CIAWITA
12	RECHU ALAMASTAH	SMK	FAIRFUD	800000	DUSUN CIAWITA

Gambar 6. Data excel dari operator set role

Gambar 6. Hasil data excel yang telah diimport menggunakan operator *Set role*.

4.1.2. Preprocessing / Cleaning

NO	NAMA PENERIMA	KATEGORI	NAMA PEMILIK	NOMINAL	Alamat
1	FAUZAN HANIF	SMP	AAH SATI AH	375.000	DUSUN CIAWITAU
2	ARIF RAHMAN	SMP	LULUNG	375.000	DUSUN CIAWITAU
3	ADAH	LANZIA	ADAH	600.000	DUSUN CIAWITAU
4	NISA KHOERUNISA	SMA	ADE HERAWATI	500.000	DUSUN CIAWITAU
5	ANNISA	SMP	ADE HERAWATI	375.000	DUSUN CIAWITAU
6	ANAH	LANZIA	ANAH	600.000	DUSUN CIAWITAU
7	KOKOM KOMALASARI	IBU HAMIL	KOKOM KOMALASARI	750.000	DUSUN CIAWITAU
8	MAESAROH	LANZIA	MAESAROH	600.000	DUSUN CIAWITAU
9	CUCU ROKAYAH	IBU HAMIL	CUCU ROKAYAH	750.000	DUSUN CIAWITAU
10	LIA APRILIJANI	SD	IMAS	225.000	DUSUN CIAWITAU
11	MUHAMMAD ABDAN	BALITA	IMAS	750.000	DUSUN CIAWITAU
12	RIDHO ALAMSYAH	SMA	NUNUNG	500.000	DUSUN CIAWITAU
13	NONONG	SMP	NUNUNG	375.000	DUSUN CIAWITAU
14	JUBAEDAH	LANZIA	JUBAEDAH	600.000	DUSUN CIAWITAU
15	ADE HUDAYA	LANZIA	ADE HUDAYA	600.000	DUSUN CIAWITAU
16	MUHAMMAD HILMI	SD	ADE HERAWATI	225.000	DUSUN CIAWITAU
17	ANAH	LANZIA	ANAH	600.000	DUSUN CIAWITAU
18	GUNAWAN	SMA	KOKOM KOMARIAH	500.000	DUSUN CIAWITAU
19	AGUS GUSTIANA	SMA	SANTI	500.000	DUSUN CIAWITAU
20	CITRA	SD	SANTI	225.000	DUSUN CIAWITAU

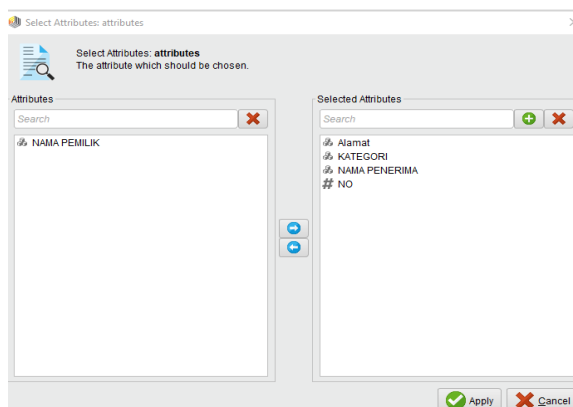
Gambar 7. Data sebelum preprocessing

Pada proses tahapan data preprocessing dilakukan pemisahan atribut dari data yang digunakan dan tidak digunakan sehingga tidak adanya missing value.

Name	Type	Missing	Statistics	Filter (5/5 attributes)
NO	Integer	0	Min: 1, Max: 330, Average: 165.500	
cluster	Nominal	0	cluster_0 (46), cluster_3 (117), cluster_1 (66), [2 more]	
NAMA PENERIMA	Real	0	Min: -1.301, Max: 1.336, Average: -0.010	
KATEGORI	Real	0	Min: -0.674, Max: 3.373, Average: 0.642	
Alamat	Real	0	Min: -0.674, Max: 1.349, Average: 0.139	

Gambar 8. Hasil proses preprocessing

Pada gambar 7 merupakan data sebelum preprocessing Sementara gambar 8 merupakan proses data preprocessing / cleaning dimana atribut yang tidak digunakan akan dihilangkan, agar data yang dipergunakan dalam proses data mining lebih efisien.



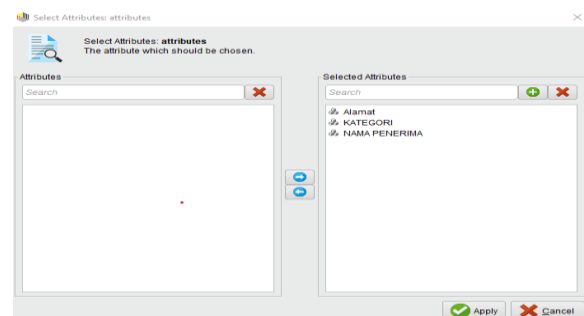
Gambar 9. Hasil data select attributes

Pada gambar 9 proses data preprocessing atau cleaning dilakukan dengan menggunakan software RapidMiner versi 9.9 didapatkan hasil bahwa tidak ada missing value pada data, baik pada atribut maupun id. Setelah didapatkan proses data cleaning dilakukan

proses seleksi dengan ekspresi reguler atau hanya memilih atribut tanpa nilai yang hilang untuk memudahkan pemilihan atribut, hasil atribut yang dipilih akan dikirim ke port output dan sisanya dihapus dari *ExampleSet*.

4.1.3. Data Transformation

Pada tahap data transformation data dilakukan perubahan data dengan menjumlahkan variabel pada jenis kategori dari setiap dusun di Desa Gereba. Maka dari 7 atribut di transformasikan menjadi 3 atribut yaitu alamat, nama penerima dan kategori. Berikut merupakan gambar proses data transformation menggunakan operator nominal to numerical.



Gambar 7. Hasil nominal to numerical

Row No.	NO	NAMA PENE...	KATEGORI	Alamat
1	1	0	0	0
2	2	1	0	0
3	3	2	1	0
4	4	3	2	0
5	5	4	0	0
6	6	5	1	0

Gambar 8. Hasil data transformation

4.1.4. Data Mining



Gambar 9. Model algoritma k-means

Tahapan pada data mining dilakukan menggunakan teknik klasterisasi dengan algoritma yang digunakan yaitu K-means. Pada operator clustering ini diambil objek dari port input dan mengirimkan salinannya ke port output, setiap port yang terhubung akan membuat salinan yang tidak terikat, oleh karena itu ketika mengubah suatu salinan tidak berpengaruh pada salinan yang lain. Gambar 12 merupakan hasil dari clustering tersebut dapat dihubungkan dengan performance yaitu untuk mengetahui suatu model algoritma K-means dengan hasil pemodelan berupa pola informasi pada salinan yang lain.

Adapun tahapan-tahapan yang dilakukan dalam proses pemodelan algoritma *K-means* yaitu :

1. Menentukan jumlah *cluster*

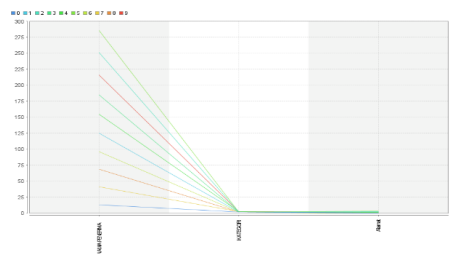
Untuk menentukan jumlah *cluster* pada *RapidMiner* yaitu menggunakan operator *clustering*, operator ini merupakan operator utama untuk mengetahui hasil dari pengklasteran tersebut. Algoritma *K-Means* digunakan untuk menentukan satu set *k cluster* dan menetapkan setiap contoh ke satu *cluster* yang tepat.

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5	cluster_6	cluster_7	cluster_8	cluster_9
NAMA PEL.	12.848	124.967	251.595	185.242	154.412	288	96.500	41.559	68.552	216.243
KATEGORI	1.606	2.100	1.730	1.606	2.235	1.914	1.750	2.206	1.931	2.378
Alamat	0.303	0.067	2.432	1.879	1.059	3	0	0.353	0.172	2.054

Gambar 10. Percobaan 10 *cluster*

Untuk menentukan jumlah *cluster* dilakukan percobaan 10 *cluster* yaitu C0 sampai C9 dengan atribut Nama Penerima, Kategori, Alamat. Berikut ini merupakan nilai *cluster* per atribut dengan jumlah percobaan 10 *cluster*

- Menentukan titik *centroid* secara acak
Titik *centroid* dipilih secara acak melalui serangkaian percobaan *cluster* di *RapidMiner*. Dari 10 percobaan tersebut, titik pusat klaster terkecil dipilih untuk memperoleh hasil yang optimal.
- Menghitung jarak data ke *centroid*
Setelah memasukkan dataset dan algoritma *K-Means* pada *RapidMiner* maka proses data dengan klik *run* untuk melihat jarak data ke *centroid*.



Gambar 11. Jarak data ke *centroid*

- Perbarui nilai titik *centroid*
Setelah beberapa percobaan, titik pusat klaster tidak mengalami perubahan dalam klaster tersebut. Jika terjadi perubahan pada titik pusat klaster, iterasi akan terus dilakukan. Namun, jika tidak ada perubahan, proses iterasi dihentikan dan kelompok-kelompok yang telah terbentuk akan diperoleh.

4.1.5. Data Mining

Untuk mengukur performa dari operator *K-Means* digunakan operator *Performance* untuk evaluasi kinerja metode *clustering* berbasis *centroid*. Operator *Cluster Distance Performance* mengambil model *cluster centroid* ini dan *clustered set* sebagai input dan mengevaluasi kinerja model berdasarkan *centroid cluster*.

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: 79.185
Avg. within centroid distance_cluster_0: 71.003
Avg. within centroid distance_cluster_1: 69.851
Avg. within centroid distance_cluster_2: 104.359
Avg. within centroid distance_cluster_3: 80.287
Avg. within centroid distance_cluster_4: 71.713
Avg. within centroid distance_cluster_5: 103.907
Avg. within centroid distance_cluster_6: 67.723
Avg. within centroid distance_cluster_7: 60.285
Avg. within centroid distance_cluster_8: 61.902
Avg. within centroid distance_cluster_9: 90.957
Davies Bouldin: 0.514
```

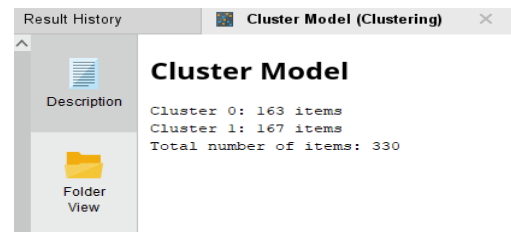
Gambar 12. Hasil *performance vector*

4.2. Analisis Hasil Penerapan Algoritma *K Means*

Berikut merupakan hasil penerapan algoritma *K-Means*:

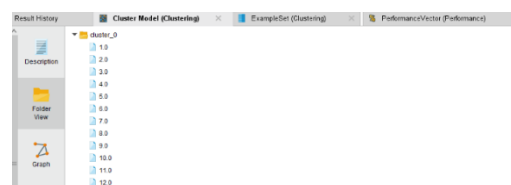
4.2.1. Hasil Analisis Penerima Bantuan PKH Dalam Bentuk *Cluster Model*

Tiap *cluster* memiliki anggota yang berbeda, berdasarkan kemiripan data satu sama lainnya. Pada penelitian ini pengelompokan dengan algoritma *K-Means* dibagi menjadi 2 *cluster* pada *cluster 0* berjumlah 163 *items*, *cluster 1* berjumlah 167 *items*.



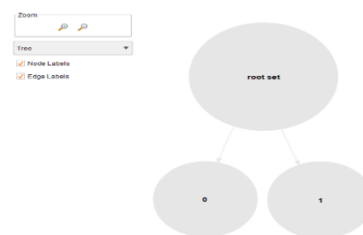
Gambar 13. *Cluster model*

Gambar 16 menampilkan hasil dari proses pengelompokan data penerima bantuan PKH menggunakan algoritma *K-Means*.



Gambar 14. *Folder View*

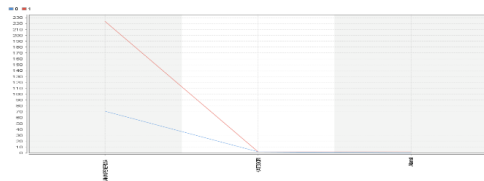
Gambar 17 menyajikan *cluster* dalam bentuk folder digunakan menu *folder view* jika di hitung jumlah anggotanya akan sama dengan jumlah yang tertera pada *description*.



Gambar 15. *Graph*

Gambar 18 merupakan ilustrasi *graph* dalam bentuk *tree*, semakin banyak anggota maka semakin

besar bulatannya. Maka dari *graph* tersebut dapat disimpulkan bahwa *cluster* 1 paling besar bulatannya dikarenakan memiliki jumlah anggota paling banyak.

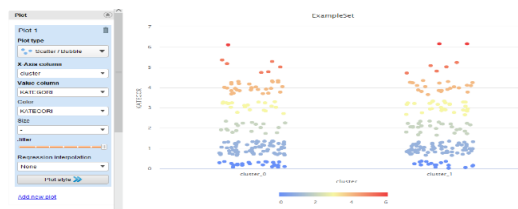


Gambar 16. Plot

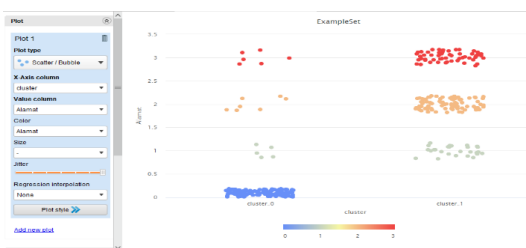
Gambar 19 menampilkan visualisasi *centroid* dalam bentuk grafis menggunakan menu *Plot* untuk memudahkan dalam membaca data.

4.2.2. Hasil Analisis Penerima Bantuan PKH Dalam Bentuk Example Set

Berikut gambar *Visualizations* Penerima Bantuan PKH berdasarkan nilai atribut kategori :



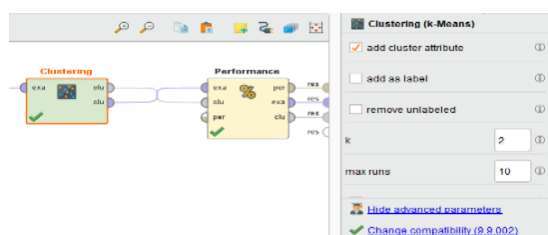
Gambar 1 Visualizations atribut kategori



Gambar 2 Visualizations atribut alamat

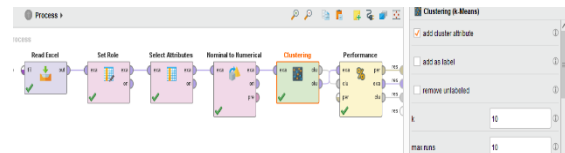
4.2.3. Menentukan Nilai Cluster Optimum dan Nilai DBI

Davies Bouldin Index atau nilai *DBI* ini bekerja untuk mengevaluasi kinerja dari proses pengklusteran yang telah dilakukan pada tahap 4.1, dari evaluasi tersebut akan menghasilkan sebuah parameter berupa *numeric*. Berikut merupakan proses pencarian nilai *DBI* :

Gambar 3 Tampilan parameter *k-means clustering*

Pada Gambar 22 merupakan proses menentukan cluster optimum dan nilai DBI dari mengelompokkan

penerima bantuan PKH menggunakan algoritma *K-Means*. Perlu dilakukan uji coba n nilai cluster 2-10, dari setiap cluster yang dibuat dan dievaluasi dengan operator *Cluster distance performance* untuk mengetahui nilai DBI tiap cluster 2-10. Pada penelitian ini menggunakan 3 metode, yaitu tanpa menggunakan operator *Normalize*, *Z-Transformation* dan *Range Transformation*.

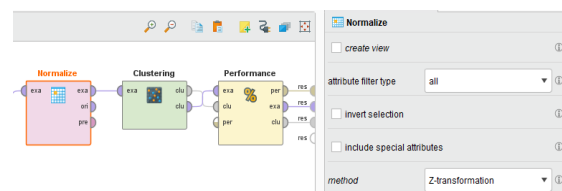
Gambar 4 Pencarian nilai DBI tanpa menggunakan operator *normalize*

Pada gambar 23 adalah hasil pencarian nilai DBI tanpa menggunakan operator *normalize* dengan percobaan K2 sampai K10. Berikut merupakan tabel pencarian nilai DBI tanpa menggunakan operator *normalize*:

Tabel 1. Pencarian nilai DBI tanpa menggunakan operator *normalize*

Cluster (K)	Nilai DBI
2	0.500
3	0.500
4	0.505
5	0.504
6	0.516
7	0.520
8	0.522
9	0.516
10	0.514

Pencarian yang kedua menggunakan operator *Normalize* dengan *method Z-Transformation*. Normalisasi ini berfungsi untuk menskalakan nilai agar sesuai dengan rentang tertentu serta untuk membandingkan atribut yang ukurannya bervariasi. *Z-Transformation* ini juga disebut normalisasi statistik untuk mengurangi rata-rata data dari semua nilai, dan mempertahankan distribusi data asli agar tidak terlalu terpengaruh oleh outlier.

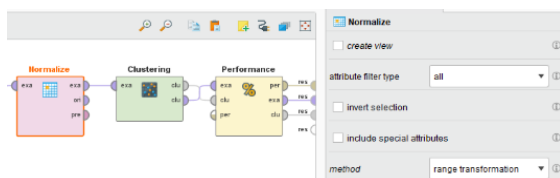
Gambar 5 Pencarian nilai DBI menggunakan operator *Z-Transformation*

Pada gambar 24 adalah hasil pencarian nilai DBI menggunakan operator *normalize* dengan metode *Z-Transformation*. Berikut merupakan tabel pencarian nilai DBI menggunakan metode *Z-Transformation* dengan percobaan K2 sampai K10:

Tabel 2. Hasil pencarian nilai DBI metode z-transformation

Cluster (K)	Nilai DBI
2	0.922
3	0.758
4	0.828
5	0.863
6	0.694
7	0.794
8	0.892
9	0.941
10	0.882

Percobaan yang ketiga yaitu menggunakan metode *Range Transformation*. Metode ini menormalkan semua nilai Atribut ke rentang nilai tertentu dan semua nilai lainnya diskalakan, sehingga sesuai dengan rentang yang diberikan. Metode *Range Transformation* ini juga dapat dipengaruhi oleh outlier, karena batasnya bergerak ke arah outlier. Berikut merupakan gambar pencarian nilai DBI menggunakan operator *Normalize* dengan metode *Range Transformation*:



Gambar.25. Pencarian nilai DBI metode range transformation

Pada gambar 25 adalah hasil pencarian nilai DBI menggunakan operator *normalize* dengan metode *Range Transformation*.

Tabel 3. Hasil pencarian nilai DBI metode z-transformation

Cluster (K)	Nilai DBI
2	0.729
3	0.848
4	0.982
5	0.853
6	0.853
7	0.908
8	0.793
9	0.626
10	0.681

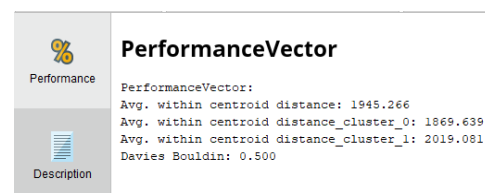
Berdasarkan ketiga pengujian yang telah dilakukan, maka hasil rekapitulasi perbandingan nilai DBI dapat dilihat pada tabel 4 berikut ini:

Tabel 3. Rekapitulasi 3 metode

K	Nilai DBI		
	Non Normalize	Z Transformation	Range Transformation
2	0.500	0.922	0.729
3	0.500	0.758	0.848
4	0.505	0.828	0.982

K	Nilai DBI		
	Non Normalize	Z Transformation	Range Transformation
5	0.504	0.863	0.853
6	0.516	0.694	0.853
7	0.520	0.794	0.908
8	0.522	0.892	0.793
9	0.516	0.941	0.626
10	0.514	0.882	0.681

Dari hasil pencarian 3 metode tersebut maka *cluster* terbaik ada di *cluster* 2 tanpa menggunakan operator *normalize*. Berikut merupakan hasil *Performance Vector* pada *cluster* 2 :



Gambar 26. Hasil performance vector pada cluster

5. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang telah dilakukan di Desa Gereba dengan menggunakan 330 data maka dapat diuraikan kesimpulan sebagai berikut: Pengelompokan dengan algoritma K-Means dibagi menjadi 2 *cluster* pada *cluster* 0 berjumlah 163 *items*, *cluster* 1 berjumlah 167 *items*. Setelah dilakukan proses *clustering* maka dapat diketahui *cluster* yang menjadi prioritas lebih di Desa Gereba adalah berada pada *cluster* 1 yaitu Dusun Ciawitali. Cluster optimum optimum dari *DBI* didapat pada *cluster* 2 tanpa menggunakan operator *Normalize* dengan nilai *DBI* yaitu 0.500. Disarankan untuk pengembangan penelitian selanjutnya dapat dilakukan proses implementasi *data mining* dapat dilakukan menggunakan metode yang berbeda seperti metode *Fpgrowth* yang berguna untuk menemukan himpunan data yang sering muncul atau *Association* atau metode *Decission Tree* yang berguna untuk pengambilan keputusan ataupun.

DAFTAR PUSTAKA

- [1] S. Saragi, M. U. Batoebara, and N. A. Arma, "Analisis Pelaksanaan Program Keluarga Harapan (Pkh) Di Desa Kota Rantang Kecamatan Hamparan Perak," *Publik J. Manaj. Sumber Daya Manusia, Adm. dan Pelayanan Publik*, vol. 8, no. 1, pp. 1–10, 2021, doi: 10.37606/publik.v8i1.150.
- [2] V. Museliza, A. Afrizal, and R. Eliza, "Pengaruh Program Keluarga Harapan (Pkh) Terhadap Kesejahteraan Keluarga Penerima Manfaat Di Kecamatan Tampan Kota Pekanbaru," *J. Manaj. dan Ilmu Adm. Publik*.
- [3] P. Mai Sarah Tarigan, J. Tata Hardinata, H. Qurniawan, M. Safii, and R. Winanjaya, "Implementasi Data Mining Menggunakan

- Algoritma Apriori Dalam Menentukan Persediaan Barang (Studi Kasus: Toko Sinar Harahap),” *Just IT J. Sist. Informasi, Teknol. Inf. dan Komput*
- [4] M. R. Nugroho, I. E. Hendrawan, and P. P. Purwantoro, “Penerapan Algoritma K-Means Untuk Klasterisasi Data Obat Pada Rumah Sakit ASRI,” *Nuansa Inform.*
- [5] S. N. Arofah and F. Marisa, “Penerapan Data Mining untuk Mengetahui Minat Siswa pada Pelajaran Matematika menggunakan Metode K-Means Clustering,” *JOINTECS (Journal Inf. Technol. Comput. Sci)*
- [6] M. Rafi Nahjan, Nono Heryana, and Apriade Voutama, “Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell,” *JATI (Jurnal Mhs. Tek. Inform)*
- [7] Y. Apriyanti, E. Lorita, and Y. Yusuarsono, “Kualitas Pelayanan Kesehatan Di Pusat Kesehatan Masyarakat Kembang Seri Kecamatan Talang Empat Kabupaten Bengkulu Tengah,” *Prof. J. Komun. dan Adm. Publik*, vol. 6, no. 1, 2019, doi: 10.37676/professional.v6i1.839.
- [8] S. Widaningsih, “Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naïve Bayes, Knn Dan Svm,” *J. Tekno Insentif*, vol. 13, no. 1, pp. 16–25, 2019, doi: 10.36787/jti.v13i1.78.
- [9] K. F. Mauladi and P. H. Susilo, “Klasterisasi Virus Covid-19 Di Wilayah Kabupaten Lamongan Dengan Metode K-Means Clustering,” *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*
- [10] I. K. Juni Arta, G. Indrawan, and G. R. Dantes, “Data Mining Rekomendasi Calon Mahasiswa Berprestasi Di Stmik Denpasar Menggunakan Metode Technique for Others Reference By Similarity To Ideal Solution,”