

PENERAPAN DATA MINING CLUSTERING MENGGUNAKAN ALGORITMA X-MEANS PADA DATA PENERIMA BANTUAN PROGRAM KELUARGA HARAPAN DI DESA GEBANG KULON KABUPATEN CIREBON

Winayah, Rudi Kurniawan, Yudhistira Arie Wijaya

Teknik Informatika, STMIK IKMI CIREBON

Jalan Perjuangan No.10B, Karyamulya, Kecamatan Kesambi, Kota Cirebon, Jawa Barat 45135

Winayah2303@gmail.com

ABSTRAK

Program Keluarga Harapan yang diartikan sebagai bantuan sosial kepada keluarga tidak mampu, menghadapi kerumitan dalam pengolahan data pada penerima PKH di Desa Gebang Kulon. Kesulitan ini melibatkan penentuan prioritas penerima bantuan yang dipengaruhi oleh faktor-faktor seperti pengangguran, penghasilan rata-rata, dan jumlah anggota keluarga. Tujuan penelitian ini untuk menentukan jumlah cluster yang optimal, mengukur besarnya Davies Bouldin index, dan mengidentifikasi parameter x-means yang efektif. Dalam penerapan data mining metode yang digunakan adalah Knowledge Discovery in Databases (KDD) diterapkan dalam tahapan selection, cleaning, transformation, dan evaluation untuk memproses data penelitian. Penerapan algoritma x-means menggunakan RapidMiner versi 10.3 pada data tahun 2023 menghasilkan 3 cluster, cluster 0 terdiri 1217 item, cluster 1 terdiri 3 item dan cluster 2 terdiri 4 item, dimana cluster ini mencerminkan karakteristik kelompok penerima bantuan PKH. Nilai Davies Bouldin Index sebesar -0,005 dan parameter yang digunakan yaitu Numerical Measuremet. Diharapkan bahwa hasil penelitian ini dapat memberikan kontribusi positif dengan membantu Desa Gebang Kulon dalam meningkatkan proses penyaluran bantuan PKH, sehingga pemberian bantuan menjadi lebih optimal dan akurat sesuai masyarakat yang membutuhkan.

Kata kunci: Data Mining, Clustering, X-Means, Davies Bouldin Index (DBI), dan Program Keluarga Harapan

1. PENDAHULUAN

Kemiskinan merupakan masalah serius dalam proses pembangunan, selain kegiatan pembangunan ada berbagai program pengentasan kemiskinan yang telah diupayakan. Program bantuan sosial yang telah pemerintah jalankan saat ini, salah satunya Program Keluarga Harapan adalah program bantuan untuk keluarga berpenghasilan rendah, permasalahannya, terkadang proses penyaluran bantuan ini masih dianggap tidak merata atau tepat sasaran.

Data mining merupakan bagian integral dari proses *Knowledge Discovery in Databases* (KDD), di mana langkah-langkahnya melibatkan ekstraksi pengetahuan dari data. Proses KDD mencakup pemelihan data, preprocessing, transformasi, data mining, interpretasi atau evaluasi, dan pembentukan pengetahuan baru. Clustering adalah salah satu teknik data mining yang digunakan untuk mengelompokkan data berdasarkan kesamaan fitur atau karakteristik tertentu. Dalam melakukan clusterisasi, beberapa hal yang perlu di pertimbangkan jenis data, tujuan clusterisasi, metode clusterisasi, preprocessing data dan evaluasi.

Algoritma X-Means adalah variasi dari algoritma K-Means yang dikembangkan untuk mengatasi kelemahan dalam menentukan jumlah cluster secara manual. X-Means menggunakan pendekatan yang lebih dinamis untuk menentukan jumlah cluster, tidak memerlukan input K dari pengguna, dan bekerja secara interative. Meskipun tidak selalu lebih cepat dari K-Means, X-Means dapat lebih efisien dalam beberapa kasus karena menghindari kebutuhan untuk mengulangi eksperimen dengan jumlah cluster yang berbeda.

Penelitian ini bertujuan untuk menentukan jumlah cluster dan mengukur besarnya *Davies Bouldin Index*. Sehingga penerapan Data Mining dengan metode Clustering X-Means, untuk memudahkan pengelolaan data, efektifitas pengelompokan dapat diukur dari tingkat homogenitas dalam suatu kelompok dan tingkat perbedaan antar kelompok. Semakin tinggi tingkat homogenitas dan perbedaan, semakin baik hasil pengelompokan. Algoritma yang mampu mencapai pengelompokan ini adalah x-means. Pada X-Means, salahsatu kelemahannya pada penentuan matriks jarak. Matriks jarak menjadi faktor kursial yang bergantung pada dataset yang digunakan oleh algoritma X-Means. Kualitas nilai matriks jarak yang dihasilkan akan berdampak pada kinerja algoritma tersebut.

Dari latar belakang yang telah diuraikan, penelitian ini akan membahas mengenai “Penerapan Data Mining Clustering Menggunakan Algoritma X-Means Pada Data Penerima Bantuan Program Keluarga Harapan” Harapannya, temuan dari penelitian ini dapat memberikan dukungan kepada pihak Desa Gebang Kulon dalam meningkatkan efektivitas penyaluran bantuan Program Keluarga Harapan (PKH). Dengan demikian, diharapkan pemberian bantuan menjadi lebih optimal dan terarah kepada masyarakat yang membutuhkan.

2. LITERATURE REVIEW TINJAUAN PUSTAKA

Pencarian literatur yang relevan dengan topik penelitian. Langkah ini dapat memberikan gambaran mengenai Penerapan Data Mining Clustering

Menggunakan Algoritma X-Means Proses pencarian dan pengumpulan artikel dengan menggunakan tools publish or perish, yang diambil dari database akademik yaitu *Google Scholar*, *Shinta* dan *Scopus* dengan kata kunci “Data Mining” dan “Clustering X-Means”.

[1] dengan judul “Analisa Penerapan Metode Clustering X-Means Dalam Pengelompokan Penjualan Barang” membahas mengenai pengetahuan tingkat penjualan produk pada sebuah perusahaan untuk memberikan rekomendasi pada pemilik perusahaan tentang produk mana yang prioritas diberi perhatian demi kemajuan perusahaan tersebut. Metode yang digunakan dalam penelitian ini yaitu algoritma clustering x-means.

[2] dengan judul “Analisis Rumah Tidak Layak Huni Menggunakan Algoritma X-Means Analyzing Uninhabitable Houses Using X-Means Algorith” membahas mengenai pengelompokan rumah tidak layak huni dengan melihat struktur bangunan rumah dan fasilitas rumah yang dimiliki setiap rumah tangga di 4 wilayah yaitu: Batu layang, siantan hilir, siantan tengah dan siantan hulu. Salah satu metode yang dapat digunakan untuk pengelompokan rumah tidak layak huni adalah metode X-Means clustering. Dalam memetakan atau mengelompokkan rumah tidak layak huni menggunakan algoritma x-means menjadi beberapa cluster serta untuk mengetahui nilai akurasi terbaik dari hasil uji Davies Bouldin Index (DBI).

[3] dengan judul “Data Mining Manhattan Distance dan Euclidean Distance pada Algoritma X-Means dalam Klasifikasi Minat dan Bakat Siswa” membahas mengenai Data mining yang dibangun dapat membantu dan memberikan kemudahan dalam proses klasifikasi minat dan bakat siswa. Setting variabel penilaian dilakukan melalui sistem. Jadi, jika ada perubahan nilai maksimal dan minimal variabel, admin tinggal mengupdate melalui sistem tanpa harus merubah programnya.

[4] dengan judul “Analisis Clustering Penduduk Miskin Di Provinsi Aceh Menggunakan Algoritma K-Means Dan X-Means” membahas mengenai perbandingan hasil analisis clustering K-Means dengan X-Means menggunakan data penduduk miskin provinsi Aceh tahun 2018 dan data Iris. Hasil pengujian clustering dengan data Iris memiliki tingkat kemurnian dan hasil pengelompokan yang baik dari masing-masing Algoritma, Berdasarkan Waktu yang dibutuhkan selama proses clustering algoritma X-Means memiliki kecepatan yang lebih baik yaitu 0.06 second. Sedangkan K-Means membutuhkan waktu selama 0.22 second. Sehingga dapat disimpulkan bahwa algoritma X-Means lebih cepat dalam melakukan pengelompokan dibandingkan algoritma K-Means.

[5] dengan judul “Penerapan Algoritma Clustering K-Means Untuk Menentukan Prioritas Penerima Bantuan Dana Sosial PKH Di Kelurahan Kampung Singkep” membahas mengenai perhitungan Clustering K-Means untuk menentukan penerima bantuan terhadap data penduduk Kelurahan Kampung

Singkep, agar dapat mengetahui bantuan apa yang sesuai dengan kondisi mereka, sehingga mempermudah tim penyeleksi memberikan bantuan sesuai dengan kriteria yang telah ditentukan apakah layak atau tidak layak menerima bantuan tersebut.

[6] dengan judul “Penerapan Data Mining Untuk Bantuan Program Keluarga Harapan Menggunakan Algoritma K-Means” membahas mengenai penggunaan Algoritma K-Means dalam proses klasterisasi data penerima PKH, maka diprediksi akan ditentukan kelompok prioritas dan nomor klaster yang paling tepat/akurat untuk data penerima PKH di Kecamatan Banjarmasin Tengah kedepannya memastikan bahwa data tidak salah arah. Algoritma K-Means digunakan dalam perangkat lunak. Diasumsikan bahwa setelah menggunakan perangkat lunak, model desain yang digunakan dalam hal ini akan selesai dan memiliki solusi terbaik. Sehingga nantinya dapat membantu Camat Banjarmasin Tengah dalam menentukan kategori prioritas untuk mendapatkan dana bantuan PKH.

[7] dengan judul “Penerapan Data Mining Untuk Menentukan Kelompok Prioritas Penerima Bantuan PKH Menggunakan Metode Clustering K-Means Pada Desa Kuala Dendang” membahas mengenai karena banyaknya data penduduk pada desa Kuala Dendang, membuat pihak yang menangani data membutuhkan waktu lama untuk menentukan warga yang berhak mendapatkan bantuan karena masih melakukan pengecekan data secara manual, yaitu dengan melihat data satu-persatu. Apalagi data yang tersedia masih ada beberapa yang belum diperbaharui atau lengkap. Atau terkadang pihak desa hanya menerima rekomendasi dari setiap RT yang ada mengenai warga-warga yang diusulkan mendapat bantuan PKH. Penentuan prioritas penerima bantuan ini dapat dilakukan dengan data mining, salah satunya dengan metode Clustering K-Means.

[8] dengan judul “Data Mining Menentukan Cluster Penerima Program Bantuan dengan Metode K-Means” membahas mengenai penggunaan data mining untuk melakukan cluster data dengan metode K-Means dan pengolahan data juga dilakukan dengan aplikasi Rapidminer. Dalam data mining, data yang dimiliki saat penelitian dilakukan pengklasifikasian data, membuat prediksi serta memperkirakan untuk mendapatkan informasi yang bermanfaat dari kumpulan data yang akan diolah.

[9] dengan judul “Pengelompokkan Tingkat Pemahaman Guru PAUD Terhadap Pembelajaran Berbasis STEAM Menggunakan Metode X-Means Clustering” membahas mengenai penggunaan algoritma X-Means clustering sebagai salah satu algoritma data mining dalam melakukan pengelompokan data tingkat pemahaman guru PAUD terhadap pembelajaran berbasis STEAM. Pengelompokkan dengan menggunakan algoritma X-Means Clustering menghasilkan tingkat pemahaman guru PAUD di Kecamatan Medan Area sebanyak 12 guru yang sangat paham, 7 guru yang paham, 7 guru

yang cukup paham, serta 4 guru yang tidak paham pembelajaran berbasis STEAM.

[10] dengan judul “Implementation of the X-Means Algorithm on Unemployment Data in West Java” membahas mengenai pengelompokan data pengangguran khususnya di Provinsi Jawa Barat dengan menerapkan algoritma X-Means berdasarkan jenjang pendidikan dan kabupaten/kota untuk melihat kenaikan dan penurunan dari pengangguran. Hasil penelitian menunjukkan bahwa terdapat fluktuasi dalam jumlah pengangguran dalam Provinsi Jawa Barat di tahun 2011 hingga 2022. Tingkat pengangguran rendah di Provinsi Jawa Barat meningkat sebesar 66,7% pada 2014-2015, stabil pada 2017-2019, sementara tingkat pengangguran tinggi turun 80% pada 2014-2015, stabil pada 2017-2019, tetapi naik 20% pada 2020-2022.

3. METODE PENELITIAN

Penelitian ini dilakukan berdasarkan tahapan-tahapan penting yang dikerjakan dengan berorientasikan kepada indikator keberhasilan dalam pengelompokan data penerima PKH yang ada di Desa Gebang Kulon sehingga dapat digunakan untuk menyelesaikan permasalahan multi objektif. Untuk dapat mencapai, indikator tersebut, maka tahapan-tahapan penelitian ini adalah sebagai berikut.

- Analisa masalah, dalam hal ini kebutuhan menganalisa permasalahan yang akan diteliti mengenai pengelompokan data masyarakat miskin.
- Analisa kebutuhan, dalam hal ini segala kebutuhan dalam meneliti baik dari jurnal, buku, literatur-literatur, tools dan software.
- Mendesain system clustering berdasarkan Algoritma X-Means.
- Membuat analisa dengan menggunakan tools Rapidminer.
- Menguji data dengan Algoritma X-Means yang dibuat.
- Membuat laporan dan menyimpulkan hasil penelitian

Pada penelitian ini menggunakan proses *Knowledge Discovery in Databases* (KDD), Berikut tahapan penelitian yang dilakukan:

- Selection
salah satu proses dimana pengambilan informasi/pengumpulan informasi, pemilihan informasi serta pengecekan informasi. Pada tahap ini, data digunakan sebagai bahan penelitian adalah data yang diambil dari tahun 2023 lalu dipilih atribut sesuai yang dibutuhkan.
- Preprocessing
sebagian proses persiapan informasi sebelum dikerjakannya proses Data Mining. Data yang dilakukan preprocessing biasanya dicoba dengan sebagian sesi berupa *cleaning*, *reduction*, dan *integration*. Saat sebelum proses prediksi, diperlukan proses preprocessing terlebih dahulu

dengan hapus data yang tidak perlu, data yang hilang atau data missing.

- Transformasi
Pada tahap ini dilakukan proses transformasi data ke dalam bentuk format yang sesuai untuk di proses ke Data Mining. Bertujuan untuk memudahkan koordinasi data yang diproses oleh Algoritma dan alat yang digunakan dalam penelitian.
- Data Mining
proses pengolahan data berdasarkan Algoritma yang konsisten dengan alur Data Mining. Algoritma yang digunakan dalam penelitian ini adalah Algoritma X-Means Clustering.
- Evaluation
Pada tahap evaluasi digunakan untuk mengevaluasi hasil dari Algoritma X-Means Clustering untuk mengetahui kesimpulan dari proses data mining.
- Knowledge
Tahapan Knowledge ini menampilkan data yang dihasilkan dari proses Data Mining, proses ini memaparkan hasil sehingga menjadi data yang mudah untuk dipahami.

4. HASIL DAN PEMBAHASAN

4.1. Dataset

Dataset yang digunakan, dapat dilihat pada Gambar berikut ini.

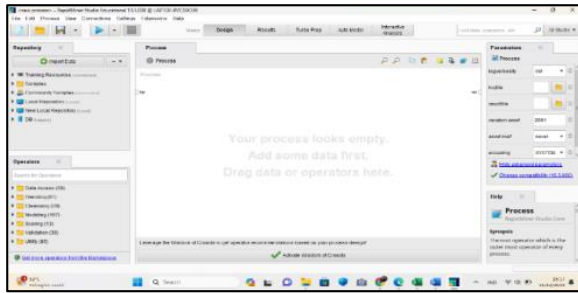
No	NPK	Nama	Jabatan	Dusun	P.T	P.W	Pekerjaan	PKH	BB	BBNT	Jumlah
1	3.2093E+15	BARUDIN	Pengurus	03	2	9	NELAYAN	YA	YA	TIDAK	2
2	3.2093E+15	ADE HASAN	Anggota	05	2	11	BELUM BEKERJA	YA	YA	TIDAK	2
3	3.2093E+15	KONVASSIH	Anggota	04	2	9	PELAJAR	YA	YA	TIDAK	2
4	3.2093E+15	ELAH	Pengurus	03	2	7	RUMAH TANGGA	YA	YA	TIDAK	2
5	3.2093E+15	DWI URYANTI	Anggota	03	2	7	PELAJAR	YA	YA	TIDAK	2
6	3.2093E+15	SULAIMAN	Anggota	04	2	8	BELUM BEKERJA	YA	YA	TIDAK	2
7	3.2093E+15	M. BENAL ARIFFIN	Anggota	06	2	14	BELUM BEKERJA	YA	YA	TIDAK	2
8	3.2093E+15	BARITA AHMAD D	Anggota	01	2	1	NELAYAN	YA	YA	TIDAK	2
9	3.2093E+15	KARNANTO	Anggota	04	1	8	BELUM BEKERJA	YA	YA	TIDAK	2
10	3.2093E+15	ANTISMI	Anggota	05	2	11	BELUM BEKERJA	YA	YA	TIDAK	2
11	3.2093E+15	WARIS	Anggota	01	2	1	BELUM BEKERJA	YA	YA	TIDAK	2
12	3.2093E+15	ELIYAH	Pengurus	05	3	11	RUMAH TANGGA	YA	YA	YA	3
13	3.2093E+15	RATU FARWAR E S	Pengurus	04	3	4	RUMAH TANGGA	YA	YA	TIDAK	2
14	3.2093E+15	TANZELA S	Anggota	06	2	12	BELUM BEKERJA	YA	YA	TIDAK	2
15	3.2093E+15	MUKA	Anggota	05	1	11	BELUM BEKERJA	YA	YA	TIDAK	2
1224	3.2093E+15	MUKA	Pengurus	02	2	6	RUMAH TANGGA	YA	YA	YA	3

Gambar 1. Dataset

Dataset yang digunakan dalam penelitian ini terdiri dari 1224 data penerima bantuan PKH di Desa Gebang Kulon. Data ini mencakup periode waktu dari tahun 2023, memberikan gambaran longitudinal tentang penerima bantuan selama satu terakhir.

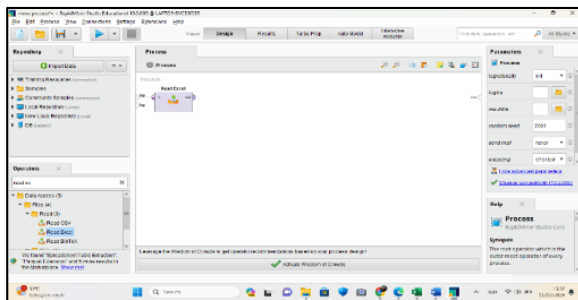
4.2. Data Mining

Proses mencari atau menggali informasi menarik dalam sebuah data dengan jumlah besar yang terpilih dengan menggunakan teknik atau metode tertentu. Penelitian ini menggunakan Algoritma X-Means untuk mengelompokkan penerima bantuan Program Keluarga Harapan (PKH) di Desa Gebang Kulon. Dalam penelitian ini pengolahan data ini menggunakan bantuan software RapidMiner. software RapidMiner bisa dilihat pada Gambar berikut ini.



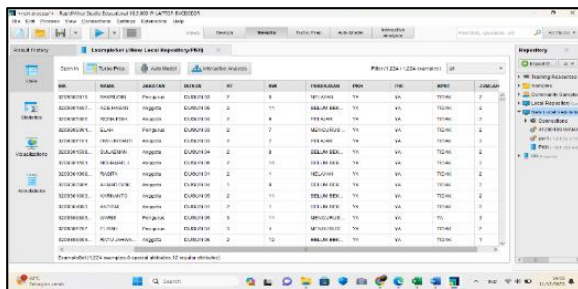
Gambar 2. Halaman Utama RapidMiner

Import data dilakukan untuk memasukkan data yang akan diuji dalam bentuk format .xls atau .xlsx. Adapun langkah untuk import data yaitu Klik kanan pada layar > Insert Operator > Data Access > Files > Read > Read Excel. Operator Read Excel yang sudah dipanggil dapat dilihat pada Gambar berikut ini.



Gambar 3. Operator Read Excel yang sudah dipanggil

Setelah operator read excel telah dipanggil, langkah selanjutnya yaitu import data untuk diproses dengan cara klik Import Configurations Wizard > klik file yang akan digunakan. Jika data telah berhasil di import dan tidak ada yang error maka klik finish untuk melanjutkan ke langkah selanjutnya. Tampilan Import Data dapat dilihat pada Gambar berikut ini.



Gambar 4. Tampilan Import Data

Pada proposes Clustering X-Means terdapat beberapa operator yang digunakan untuk menganalisis data. Operator tersebut adalah:

1. Read Excel, digunakan untuk membaca data yang akan dikelompokkan.
2. Clustering, digunakan untuk mengelompokkan ke dalam beberapa cluster.
3. Cluster Performance Distance

4.3. Selection

Untuk membaca dataset dalam bentuk file excel, menggunakan operator Read Excel seperti tampak pada Gambar berikut ini.



Gambar 5. Operator Read Excel pada Rapidminer

Setelah mengkonfigurasi operator "Read Excel" sesuai kebutuhan, dapat menggunakannya sebagai langkah awal dalam workflow RapidMiner untuk memuat data dari file Excel ke dalam platform analisis dan pemodelan data ini.

4.4. Preprocessing

Proses cleaning atau pembersihan data yang missing atau memiliki nilai yang tidak konsisten pada langkah preprocessing. Sebelum melakukan proses ini, dilakukan analisis terlebih dahulu apakah dataset memiliki nilai missing atau tidak serta memiliki data yang konsisten atau tidak. Dari hasil result dari statistik dataset seperti tampak pada Gambar dibawah ini.

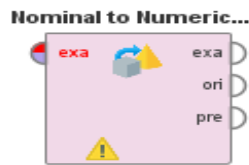
Name	Type	Missing	Statistics	Filter (14 / 14 attributes)
id	Integer	0	Min: 1 Max: 1224 Average: 612.500	
cluster	Nominal	0	Level: cluster_1 (3) cluster_0 (1217) cluster_2 (4) cluster_3 (4)	
NAMA	Nominal	0	Min: 0 Max: 1173 Average: 579.647	
JABATAN	Nominal	0	Min: 0 Max: 1 Average: 0.736	
DUSUN	Nominal	0	Min: 0 Max: 6 Average: 2.477	
PEKERJAAN	Nominal	0	Min: 0 Max: 16 Average: 3.476	
PKH	Nominal	0	Min: 0 Max: 0 Average: 0	

Gambar 6. Result dari Statistik Dataset

Diketahui bahwa tidak ada atribut yang memiliki nilai missing karena memiliki nilai 0 pada kolom missing. Untuk memeriksa konsisten atau tidak konsistennya dataset yang digunakan diperiksa per record secara langsung dan menunjukkan bahwa dataset memiliki data yang konsisten terhadap nilainya. Pada langkah Preprocessing, tidak dilakukan apa-apa karena data dianggap sudah konsisten dan tidak memiliki nilai yang missing.

4.5. Transformation

Pada langkah transformation dilakukan untuk data bertipe *polynomial* menjadi numerik agar data dapat diolah berdasarkan jarak. Untuk mengubah data operator Nominal to Numerical. Operator nominal to numerical data adalah untuk melakukan validasi sederhana dengan membagi dataset secara acak menjadi dua data terpisah yaitu data latih dan data uji. Seperti pada Gambar berikut ini.



Gambar 7. Operator Nominal to Numerical pada Rapidminer

Proses ini membantu dalam mempersiapkan data untuk analisis lebih lanjut, terutama ketika menggunakan algoritma-algoritma atau metode analisis yang memerlukan tipe data numerik. Untuk Parameter hasil data transformasi seperti pada Tabel dibawah ini.

Tabel 1. Parameter Nominal to Numerical

Parameter	Keterangan
Attribute filter type	Subset
Attribute	Select Attribute (PKH, PBI, BPNT, Dusun, Jabatan, Nama, dan Pekerjaan)
Coding type	Unique Integers
Main criterion	Davies Bouldin

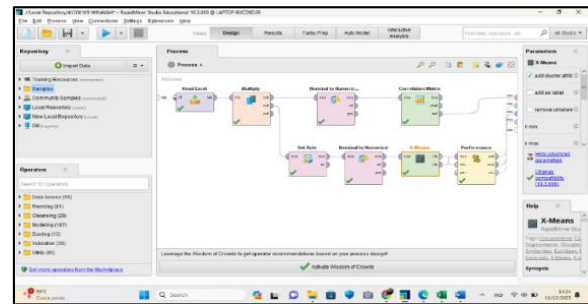
Proses ini mengubah data nominal (seperti kategori atau label) dalam atribut yang dipilih menjadi representasi numerik yang unik. Penggunaan representasi numerik ini seringkali diperlukan dalam beberapa algoritma atau analisis data yang memerlukan data dalam bentuk numerik, seperti dalam proses clustering atau analisis data mining tertentu.

Pada langkah data mining, dilakukan pengelompokan data menggunakan Algoritma X-Means sehingga menggunakan operator X-Means seperti tampak pada Gambar berikut ini.



Gambar 8. Operator X-Means pada Rapidminer

Setelah konfigurasi operator "X-Means," dapat menambahkan operator lain dalam workflow RapidMiner untuk menganalisis dan memanfaatkan hasil dari proses clustering ini. Model proses pada Rapidminer sampai dengan langkah Data Mining tampak pada Gambar berikut ini.



Gambar 9. Model proses sampai dengan langkah Data Mining

Parameter X-Means tampilannya seperti pada Tabel berikut ini.

Parameter	Keterangan
k min	3
k max	60
measure type	NumericalMeasurement
numerical measure	EuclideanDistance
clustering algorithm	KMeans

Tabel 2. Parameter X-Means

Dengan konfigurasi ini, algoritma K-Means akan dijalankan dengan mencoba membentuk kluster antara 3 hingga 60 kluster, menggunakan jarak Euclidean sebagai metrik numerik untuk mengukur kesamaan antar data.

Cluster distance performance tampilannya seperti pada Tabel berikut ini.

Tabel 3. Cluster distance performance

metode pengukuran atau kriteria utama yang digunakan untuk mengevaluasi kualitas cluster dalam konteks analisis clustering.

4.6. Evaluation

Hasil dari data mining yang diperoleh dengan menggunakan Algoritma X-Means sebagai berikut.

- 1) Hasil pengelompokan menggunakan X-Means diperoleh sebanyak 3 cluster, yaitu cluster 0 sebanyak 1217 item, cluster 1 sebanyak 3 item, dan cluster 2 sebanyak 4 item. Cluster model dapat dilihat pada Gambar berikut ini.

Cluster Model

```
Cluster 0: 1217 items
Cluster 1: 3 items
Cluster 2: 4 items
Total number of items: 1224
```

Gambar 10. Cluster Model

- 2) Nilai index Davies Boulding 0,005 untuk jumlah cluster 3. Adapun jarak rata-rata antar centroid tampak pada Gambar berikut ini.

PerformanceVector

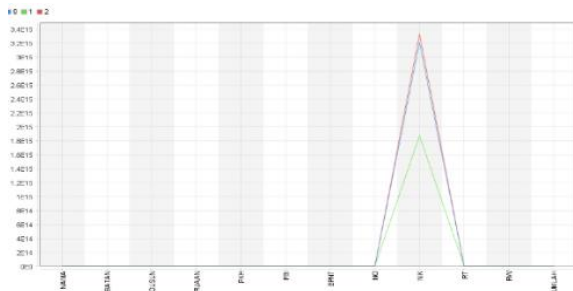
```
PerformanceVector:
Avg. within centroid distance: -66669054083314926000000.000
Avg. within centroid distance Cluster 0: -6666231804167791000000.000
Avg. within centroid distance Cluster 1: -3080851646262907400.000
Avg. within centroid distance Cluster 2: -727217974675129300000000.000
Davies Bouldin: -0.005
```

Gambar 11. Performance vector

Performance vector memberikan pemahaman komprehensif tentang seberapa baik model atau proses bekerja dan sebagai ringkasan dari evaluasi model atau analisis yang telah dilakukan.

3) Plot

Dari grafik plot diperoleh data bahwa beberapa karakteristik sebagai berikut. Grafik Plot dapat dilihat pada Gambar berikut ini.

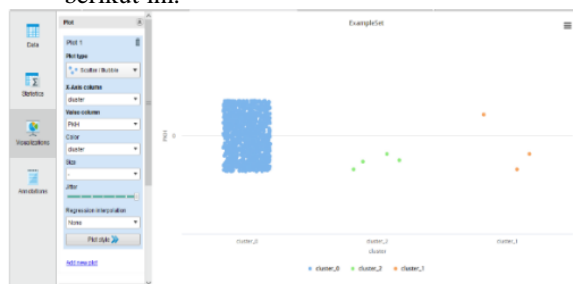


Gambar 12. Grafik Plot

Dengan menggunakan grafik plot, dapat memanfaatkan kekuatan visualisasi untuk mendukung pemahaman dan interpretasi data.

4) Visualizations

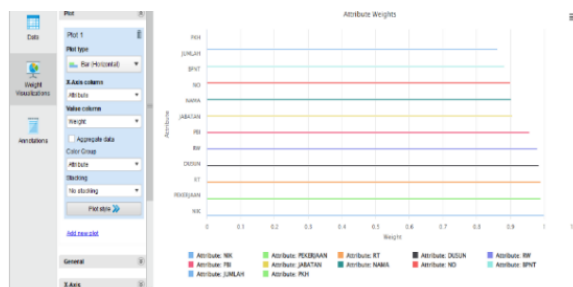
Grafik Visualizations dapat dilihat pada Gambar berikut ini.



Gambar 13. Grafik Visualizations

5) Weight Visualizations

Grafik Weight Visualizations dapat dilihat pada Gambar berikut ini.



Gambar 14. Grafik Weight Visualizations

5. KESIMPULAN DAN SARAN

Dari hasil setelah dilakukan proses Data Mining dapat disimpulkan sebagai berikut:

Hasil pengelompokan menggunakan X-Means diperoleh terbaik sebanyak 3 cluster, yaitu cluster 0 sebanyak 1217 item cluster 1 sebanyak 3 item, dan

cluster 2 sebanyak 4 item. Nilai Davies Bouldin Index yang dihasilkan dari Algoritma X-Means ini sebesar 0,005. Dalam rangka memperbaiki kekurangan serta guna penyempurnaan penelitian ini maka memberikan beberapa saran sebagai berikut: Hasil cluster yang terbentuk dapat dikembangkan dengan metode lain, seperti Neural Network, Naïve Bayes, Decision Tree, dll. Pada penelitian selanjutnya bisa menggunakan dataset dengan record lebih banyak. Penelitian ini dapat dikembangkan lagi menjadi basis pengetahuan untuk sistem pendukung Keputusan, pemetaan penduduk, dan calon penerima bantuan.

DAFTAR PUSTAKA

- [1] M. Rizqi Sulistio, N. Suarna, and O. Nurdian, "Analisa Penerapan Metode Clustering X-Means Dalam Pengelompokan Penjualan Barang," *Jurnal Teknologi Ilmu Komputer*, vol. 1, no. 2, pp. 37–42, 2023, doi: 10.56854/jtik.v1i2.49.
- [2] G. F. Rohendi, N. Suarna, and G. D. Lestari, "Analisis Rumah Tidak Layak Huni Menggunakan Algoritma X-Means," *Jurnal Janitra Informatika dan Sistem Informasi*, vol. 3, no. 1, pp. 18–29, May 2023, doi: 10.25008/janitra.v3i1.151.
- [3] M. Mughnyanti and S. Hafiz Nanda Ginting, "Data Mining Manhattan Distance dan Euclidean Distance Pada Algoritma X-Means Dalam Klasifikasi Minat dan Bakat Siswa," *remik*, vol. 7, no. 1, pp. 835–842, Jan. 2023, doi: 10.33395/remik.v7i1.12162.
- [4] B. Yusuf, R. Mahara, H. Ahmadian, S. Wahyuni, P. Studi Teknologi Informasi, and P. Sudi Pendidikan Teknologi Informasi, "Analisis Clustering Penduduk Miskin Di Provinsi Aceh Menggunakan Algoritma K-Means Dan X-Means," *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 5, no. 1, 2022.
- [5] S. Sari and J. N. Utamajaya, "Sistem Pendukung Keputusan Penerima Bantuan Langsung Tunai Dana Desa Menggunakan Metode Algoritma K-Means Clustering," 2022.
- [6] F. Yuridka, F. Teknologi Informasi, U. Islam Kalimantan MAB Banjarmasin, and A. Setiap kecamatan di Provinsi Kalimantan Selatan, "PENERAPAN DATA MINING UNTUK BANTUAN PROGRAM KELUARGA HARAPAN MENGGUNAKAN ALGORITMA K-MEANS," 2022.
- [7] S. Amaliyah and S. Rianti Agustini, "Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM) Penerapan Data Mining Untuk Menentukan Kelompok Prioritas Penerima Bantuan PKH Menggunakan Metode Clustering K-Means Pada Desa Kuala Dendang." [Online]. Available: <http://ejournal.unama.ac.id/index.php/jakakom>
- [8] R. R. A. Aria, S. Susilowati, and I. R. Rahadjeng, "Data Mining Menentukan Cluster Penerima Program Bantuan dengan Metode K-Means,"

- remik*, vol. 7, no. 1, pp. 291–300, Jan. 2023, doi: 10.33395/remik.v7i1.12030.
- [9] S. Fatimah and A. Usman, “Pengelompokkan Tingkat Pemahaman Guru PAUD Terhadap Pembelajaran Berbasis STEAM Menggunakan Metode X-Means Clustering,” 2022.
- [10] N. P. D. -, Gurbinder Singh, R. D. L. Tobing, Aldwin Simamora, J. A. T. Naibaho, and BA Wijaya, “Implementation of the X-Means Algorithm on Unemployment Data in West Java,” *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, vol. 7, no. 1, pp. 271–281, Jul. 2023, doi: 10.31289/jite.v7i1.10109.