JIKA (Jurnal Informatika) Universitas Muhammadiyah Tangerang *P ISSN : 2549-0710*
Tangerang, Januari 2024, Vol 8 No.1, pp 96-103 *E ISSN : 2722-2713*

ANALISIS DATA BALITA DENGAN PENERAPAN METODE K-MEANS CLUSTERING DI DESA WANASABA LOR

Nurul Azizah[1)], Nana Suarna[2)], Willy Prihartono[3)]

[1,2] Teknik Informatika, STMIK IKMI Cirebon, Jalan Perjuangan NO.10 B Majasem Kec. Kesambi Kota Cirebon
[3] Manajemen Informatika, STMIK IKMI Cirebon, Jalan Perjuangan NO.10 B Majasem Kec. Kesambi Kota Cirebon
Co Responden Email: na31131296@gmail.com

Article history
Received 12 Dec 2023
Revised 04 Jan 2024
Accepted 17 Jan 2024
Available online 27 Jan 2024

Keywords
Toddler,
Clustering,
K-means,
KDD

Abstract

Toddlers are children who have reached the age of 1 to under 5 years of age. The development of children under five is most important in efforts to achieve sustainable development and improve the quality of life of society. From the age of five, children experience rapid growth in language skills, creativity, social, emotional and intellectual awareness. This period became an important basis for subsequent progress. However, the health and development problems of toddlers are still a serious concern. So, careful data analysis is needed to identify trends, patterns and groups of toddlers who may want special attention. The aim of this research is to find clusters from toddler data as an effort to group 296 toddlers in Wanasaba Lor Village. Reference parameters used in clustering and calculations include gender, age, height and weight of toddlers. This research uses the methodKnowledge Discovery In Database (KDD). The results of this research include making a comparison of DBI values using the k-means algorithm by K=2 to K=10 with measure type used. You can pay attention if the cluster is close to 0, such as K=2 with a DBI value of 0.566measure type Mixed Measures, Numerical Measures, andBregman Divergences. The benefits of this research include being able to share deeper insights regarding the health situation and development of toddlers in Wanasaba Lor Village.

Riwayat
Diterima 12 Des 2023
Revisi 04 Jan 2024
Disetujui 17 Jan 2024
Terbit 27 Jan 2024

Kata Kunci
Balita,
Clustering,
K-Means,
KDD

Abstrak

Balita adalah anak yang sudah mencapai umur 1 hingga umur di bawah 5 tahun. Perkembangan anak balita paling utama saat upaya meraih pembangunan berketerusan serta menaikkan mutu hidup masyarakat. Sejak usia balita, anak mengalami pertumbuhan pesat dalam kemampuan berbahasa, kreativitas, kesadaran sosial, emosional, dan intelektual. Periode ini menjadi dasar penting untuk kemajuan berikutnya. Namun masalah kesehatan dan perkembangan balita masih menjadi perhatian serius. Sehingga, diperlukan analisis data yang cermat supaya mengidentifikasi tren, pola, dan kelompok balita yang mungkin menginginkan perhatian khusus. Tujuan atas penelitian ini untuk menemukan cluster dari data balita menjadi suatu upaya supaya mengelompokan 296 balita pada Desa Wanasaba Lor. Acuan parameter yang digunakan dalam melakukan klasterisasi dan perhitungan di antaranya seperti gender, umur, tinggi badan serta berat badan balita. Pada penelitian ini memakai metode *Knowledge Discovery In Database* (KDD). Hasil atas penelitian tersebut termasuk membuat perbandingan nilai DBI melalui algoritma k-means oleh K=2 hingga K=10 dengan *measure type* yang digunakan. Bsa diperhatikan jika cluster yang mendekati 0 seperti K=2 dengan nilai DBI 0,566 pada *measure type Mixed Measures, Numerical Measures*, dan *Bregman Divergences.* Adapun manfaat dalam penelitian tersebut termasuk dapat membagikan wawasan semakin dalam menyangkut situasi kesehatan serta perkembangan balita pada Desa Wanasaba Lor.

PENDAHULUAN

Anak usia 1 hingga di bawah lima tahun dikenal sebagai balita. (Chania, Andhini, and Jaji 2020). Masa ini memegang peran kunci dalam perkembangan fisik dan mental anak, yang nantinya akan memengaruhi keberhasilan pertumbuhan dan perkembangannya di tahap berikutnya. Dalam konteks ini, perhatian yang diberikan oleh orang tua dan pihak desa, termasuk petugas

Pusat Pelayanan Kesehatan Masyarakat (PUSKESMAS), dalam memantau kondisi balita menjadi paling penting. (Baik et al. 2022). Sehingga, mengerti aspek yang berdampak pada kesehatan dan perkembangan balita menjadi perhatian utama dalam ilmu kesehatan dan pembangunan manusia. Dalam rangka memahami faktor-faktor yang mempengaruhi kesehatan dan perkembangan balita maka perlu dilakukan pengolahan data. Pengolahan data bisa dibuat melalui memanfaatkan data mining, seperti algoritma k-means clustering. Dikarenakan pendekatan yang biasa dilakukan pada analisis cluster termasuk *k-means clustering* (Kurnia and Rinaldi 2022).

Penulis mengacu pada penelitian sebelumnya yang pertama kali dilaksanakan oleh Daniel Tunggono Saputro sebagai referensi dalam penelitian ini, Wida Pesah Sucihermayanti yang berjudul Penerapan Klasterisasi Menggunakan *K-Means* untuk Menentukan Tingkat Kesehatan Bayi dan Balita di Kabupaten Bengkulu Utara. Penelitian ini memiliki permasalahan mengenai bagaimana cara untuk mengimplementasikan algoritma k-means pada data kesehatan bayi serta balita pada Kabupaten Bengkulu Utara sehingga dapat dikelompokkan menjadi beberapa cluster yang mewakili tingkat kesehatan yang berbeda. Selain hal tersebut, tujuan atas penelitian tersebut termasuk supaya menolong Dinas Kesehatan dalam memberikan pelayanan kesehatan yang lebih baik dengan menemukan tingkat kesehatan bayi serta balita (Liesnaningsih, 2022) pada setiap kecamatan pada Kabupaten Bengkulu Utara (Saputro and Sucihermayanti 2021).

Penelitian yang kedua oleh Dona, Mi'rajul Rifqi yang berjudul Penerapan Metode *K-Means Clustering* untuk Menentukan Status Gizi Baik dan Buruk pada Balita. Penelitian ini memiliki permasalahan mengenai bagaimana cara mengatasi kekurangan gizi atau malnutrisi yang sering terjadi pada balita di Indonesia, serta kurangnya pemantauan gizi balita oleh orang tua serta aparat desa. Penelitian ini juga bertujuan untuk memberikan solusi dengan melakukan pengelompokan status gizi balita memakai metode k-means clustering (Baik et al. 2022).

Penelitian yang ketiga dilakukan oleh I Nyoman Mahayasa Adiputra melalui judul Clustering Penyakit DBD Pada Rumah Sakit Dharma Kerti Menggunakan Algoritma K-Means membahas tentang permasalahan mengenai bertambahnya jumlah pasien di Puskesmas Cigugur Tengah sepanjang harinya, makanya banyak data belum bisa dibahas semakin dalam serta datanya sekedar dipakai menjadi arsip saja. Oleh karena itu, diperlukan pengolahan data ini untuk mengelompokkan penyakit pasien berdasarkan tingkat keparahan, membedakan antara penyakit akut serta bukan akut dengan memakai teknik data mining. Metode clustering, khususnya melalui algoritma k-means serta algoritma k-medoids menjadi pembanding (Sulastri et al., 2021), diterapkan. Tujuannya adalah memberikan kontribusi bagi Puskesmas Cigugur Tengah dalam memahami jenis penyakit yang paling umum dihadapi oleh pasien, sehingga dapat mendukung pihak pemerintah, terutama Dinas Kesehatan, saat perencanaan penyuluhan kesehatan kepada masyarakat sekitar (Adiputra 2021).

Penelitian selanjutnya dikemukakan oleh Candra Adi Rahmat, Hilda Permatasari, Errissya Rasywir, Yovi Pratama yang berjudul Penerapan *K-Means* Untuk *Clustering* Kondisi Gizi Balita Pada Posyandu. Penelitian ini memiliki permasalahan mengenai bagaimana cara mengelompokan serta memilih keadaan gizi balita yang ditemukan pada Posyandu Desa Telang memakai metode k-means clustering. Penelitian ini juga berguna supaya memberlakukan metode k-means clustering saat mengelompokkan kondisi gizi balita untuk posyandu Desa Telang, serta membagikan data semakin pas serta akurat bagi petugas posyandu dalam menentukan kondisi gizi balita (Rahmat, Permatasari, and ... 2023).

Tujuan atas penelitian tersebut supaya menemukan cluster dari data balita menjadi suatu upaya supaya mengelompokan 296 balita di Desa Wanasaba Lor. Dalam pengelompokan data diperlukan analisis dan perhitungan yang cermat sehingga didapatkan kelompok balita berdasarkan tingkat kemiripan masing-masing. Acuan parameter yang digunakan dalam melakukan klasterisasi dan perhitungan di antaranya nama balita,

gender, umur, berat badan, serta tinggi badan balita.

Didalam penelitian ini penulis memanfaatkan suatu analisis dan perhitungan data yaitu data mining *algoritma k-means clustering. Clustering* termasuk suatu metode *data mining* yang termasuk pada golongan *unsupervised learning* (Prakoso et al. 2023).

Data mining merupakan suatu metode supaya menjumpai data baru yang bermanfaat bagi kumpulan data dalam jumlah yang besar untuk menolong saat pengambilan keputusan (Irawan 2019).

Clustering termasuk langkah pengelompokan titik-titik data ke dalam dua kelompok atau lebih, yang tujuannya supaya titik-titik data yang berada pada kelompok yang sama memperoleh kesamaan semakin tinggi satu sama lain daripada dengan titik-titik data pada kelompok yang tidak sama(Adiya and Desnelita 2019). Clustering memisahkan kumpulan data atas banyak kelompok dari pola yang sama (Muhima et al. 2023). Metode clustering non-hirarki berbasis jarak, seperti *algoritma k-means*, digunakan untuk memisahkan data sebagai cluster. Algoritma ini efektif untuk atribut numerik (Tinendung and Zufria 2023). Dalam metode ini, data mllaui karakteristik serupa dikelompokkan bersama pada sebuah cluster.(Chusyairi and Ramadar Noor Saputra 2019).

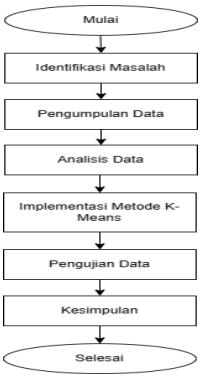
Kemudian k-means clustering memodelkan dataset yang ada dan menguji hasil *clustering* menggunakan *Cluster Distance Performance* (Afiasari, Suarna, and Rahaningsi 2023). Penelitian ini mengadopsi metode *Knowledge Discovery in Database* (KDD). KDD termasuk suatu proses komputasi yang melibatkan penerapan algoritma-algoritma matematika supaya mengekstraksi data serta menghitung probabilitas tindakan yang mungkin terjadi di masa mendatang(Muttaqin and Defriani 2020). Dalam melakukan penentuan jumlah cluster terbaik maka digunakan David Boulden Index (DBI). Davies Bouldin Index (DBI) termasuk ukuran supaya menilai kinerja clustering (Anam, Sudrajat, and Kurnia 2022).

Sebagai alat bantu dalam penelitian ini, digunakan platform perangkat lunak ilmu data bernama RapidMiner. RapidMiner menyajikan lingkungan terintegrasi supaya mempersiapkan data, melakukan pembelajaran mesin, mengeksplorasi penambangan teks, serta menganalisis prediksi. Platform ini mendukung berbagai tahapan pada proses pembelajaran, mencakup persiapan data, visualisasi hasil, validasi model, serta optimasi(Saputro and Sucihermayanti 2021).

METODE PENELITIAN

A. Metode Penelitian

Penelitian ini memberlakukan metode pendekatan kuantitatif, yang menggunakan angka sebagai sarana evaluasi untuk menganalisis informasi yang ingin diketahui. Dalam menjalankan metode penelitian, terdapat serangkaian tahapan yang harus dilalui. Tahapan-tahapan ini merupakan langkah-langkah yang perlu dijalankan sesuai dengan fokus permasalahan, agar tetap sesuai dengan batasan masalah yang telah ditetapkan dan memastikan kelancaran jalannya penelitian. Rinciannya mengenai langkah-langkah pelaksanaan penelitian dapat ditemukan dalam Gambar 1 seperti:



Gambar 1. Tahapan Penelitian

Seperti deskripsi atas tahapan penelitian lewat Gambar 1 di atas:

1) Identifikasi Masalah
   Menganalisis permasalahan yang berkaitan dengan data balita dengan menetapkan kriteria-kriteria yang relevan untuk balita.
2) Pengumpulan Data
   Mengumpulkan data yang bersangkutan pada topik penelitian dari beragam sumber yang relevan supaya mendukung penelitian ini.
3) Analisis Data
   Data yang dipakai untuk penelitian ini bermula oleh arsip posyandu Desa Wanasaba Lor terkait kondisi balita.

Setelah mendapatkan data yang diperlukan, dilakukan penentuan variabel yang nanti dipakai pada penelitian. Parameter acuan yang dipakai melibatkan usia, tinggi badan, serta berat badan balita.

4) Implementasi Metode K-Means
Proses pengolahan data dilakukan dengan menerapkan metode data mining k-means. Tujuan utamanya adalah untuk menemukan data yang bisa dipakai menjadi dasar pengambilan keputusan saat mengatasi masalah yang ada.
5) Pengujian Data dengan RapidMiner
Pengujian data dilaksanakan dengan menggunakan alat bantu *RapidMiner* untuk mendapatkan cluster mengikuti metode *k-means*.
6) Kesimpulan
Dapat mengelompokkan data balita berdasarkan kriteria tertentu.

B. Sumber Data

Sumber data dalam Analisis Data Balita Dengan Penerapan Metode K-Means Clustering ini diperoleh dari 7 Pusat Pelayanan Terpadu (POSYANDU) yang ada di wilayah Desa Wanasaba Lor tahun 2022-2023. Atribut dari data tersebut adalah nama balita, gender, umur, berat badan dan tinggi badan (Andini et al., 2023).

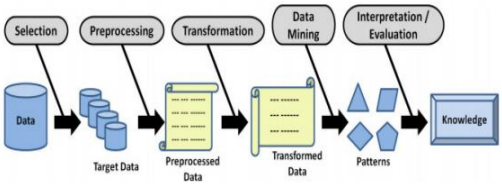
C. Teknik Pengumpulan Data

Metode pengumpulan data yang dibuat pada penelitian ini menggunakan beberapa metode di antaranya:

1. Observasi
Datang ke tempat POSYANDU yang ada di Desa Wanasaba Lor dan berinteraksi langsung dengan petugas untuk meminta data balita sehingga didapat informasi mengenai nama balita, gender, umur, berat badan serta tinggi badan balita yang relevan.
2. Literature Review
Dalam metode ini, penulis melakukan tinjauan pustaka dengan mencari berbagai sumber referensi yang relevan pada permasalahan yang nanti dibahas.

D. Teknik Analisis Data

Untuk penelitian tersebut memakai algoritma k-means melalui model KDD. Dalam KDD, suatu proyek Data Mining memperoleh siklus yang dipisahkan pada 5 tahap. Tahapan tersebut bisa diperhatikan di Gambar 2:



Gambar 2. Tahapan Knowledge Discovery In Database (KDD)

1) Data Selection
Data selection termasuk tahap awal dalam proses Knowledge Discovery In Database (KDD), dimana dari sekumpulan data yang diperoleh diadakan seleksi data terlebih dahulu, data pada wujud file Microsoft Excel mellaui format xlsx yang diperoleh kemudian import data ke dalam aplikasi Rapid Miner untuk melakukan pembacaan data menggunakan operator Read Excel.
2) Preprocessing Data
Preprocessing data berguna supaya membuang duplikasi data, mengecek data yang belum tetap serta merenovasi kekeliruan untuk data, proses ini merupakan tahapan penting karena jika tidak dilakukan dapat menghambat hasil dari proses Data Mining. tetapi hasil menunjukan pada data ini tidak memiliki Missing Value atau kesalahan pada data.
3) Transformation
Tahap transformasi dilakukan dengan menggunakan satu operator dengan tujuan untuk menyiapkan data, dengan cara mengubah data yang sudah dipilih dan dikonversi menjadi bentuk yang cocok untuk proses Data Mining, proses transformasi memakai operator Nominal To Numerical.
4) Data Mining
Data mining termasuk proses supaya menemukan pola maupun data yang ada di saat data yang sudah terpilih pada saat proses transformasi. Dalam penelitian tersebut data mining yang gunakan termasuk metode k-means. Metode k-means dapat dilakukan melalui mengukur jarak titik-titik pusat. Jarak antara dua titik dapat diukur menggunakan formula jarak Euclidean atau metode pengukuran jarak lainnya. Misalnya, jika (x_1, y_1) dan (x_2, y_2) adalah koordinat dua titik dalam suatu cluster, maka jarak Euclidean d

antara keduanya dapat dihitung memakai rumus seperti:

$$d(x,y) = \sqrt{(x_1,y_1)^2 + (x_2,y_2)^2}$$

5) Interpretation / Evaluation
Melalui proses data mining dibuat proses evaluasi terhadap data yang sudah didapat sesuai dengan prosedur yang telah ditentukan, tahap ini juga menggunakan aplikasi RapidMiner supaya menghitung keahlian metode algoritma k-means mellaui uji performance hingga menghasilkan nilai Davies Bouldin Index (DBI) terbaik. DBI dihitung dengan mempertimbangkan rasio antara dua kelompok cluster yang berdekatan dan seberapa besar nilai kepadatan dalam kelompok tersebut. Skor DBI yang lebih rendah menunjukkan bahwa pembentukan cluster lebih baik, dengan kelompok yang lebih homogen dan berdekatan.

HASIL DAN PEMBAHASAN

1) *Data Selection*

Data selection merupakan tahap pertama pada proses Knowledge Discovery In Database (KDD), dimana dari sekumpulan data yang diperoleh diadakan seleksi data terlebih dahulu, data pada wujud file Microsoft Excel mellaui format xlsx yang diperoleh kemudian import data ke ke dalam aplikasi RapidMiner untuk melakukan pembacaan data memakai operator Read Excel lewat Gambar 3 seperti :

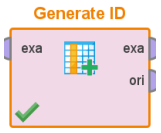
Gambar 3 Operator read excel

Hasil atas pembacaan operator Read Excel diperoleh data bisa diperhatikan lewat Tabel 1 seperti:

Tabel 1 Statistik data *read excel*

No	Uraian	Keterangan
1	Record	296
2	Special Attribute	0
3	Regular Attribute	5
4	Attribute	
	Nama Balita	Nominal, missing 0
	Jenis Kelamin	Nominal, missing 0
	Umur (Bulan)	Integer, missing 0
	Berat (Kg)	Real, missing 0
	Tinggi (Cm)	Real, missing 0

Supaya memilih ID untuk dataset maka dipakai operator *Generate ID*, misalnya Gambar 4 seperti:



Gambar 4 Operator generate id

Hasil dari pembacaan operator Generate ID didapat informasi yang bisa diperhatikan Tabel 2 seperti:

Tabel 2 Hasil operator generate id

No	Uraian	Keterangan
1	Record	296
2	Special Attribute	1
3	Regular Attribute	5
4	Attribute	
	ID	Integer, missing 0
	Nama Balita	Nominal, missing 0
	Jenis Kelamin	Nominal, missing 0
	Umur (Bulan)	Integer, missing 0
	Berat (Kg)	Real, missing 0
	Tinggi (Cm)	Real, missing 0

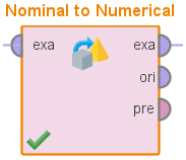
Penjelasan :
Special atribut yang dimaksud pada Tabel 2 adalah ID dengan tipe data integer.

2) *Data Preprocessing*

Langkah selanjutnya adalah melakukan preprocessing pada data yang memiliki Missing Value. Ketika sebuah atribut dalam kumpulan data yang sudah dipilih tidak memiliki nilai atau kosong maka itu diindikasikan sebagai Missing Value, tetapi hasil menunjukan pada data ini tidak memiliki Missing Value diketahui bahwa atribut memiliki nilai 0 pada tiap kolom missing, maka pada tahap preprocessing tidak perlu dilakukan.

3) *Data Transformasi*

Tahap transformasi dilakukan dengan menggunakan satu operator dengan tujuan untuk menyiapkan data, dengan cara mengubah data yang sudah dipilih dan dikonversi menjadi bentuk yang cocok untuk proses Data Mining, proses transformasi memakai operator Nominal To Numerical misalnya lewat Gambar 5 di bawah ini:



Gambar 5 Operator nominal to numerical

Operator Nominal To Numerical membantu proses data supaya dapat diproses untuk Data Mining dengan cara merubah data Polynominal menjadi Numeric, tahap ini dilakukan bertujuan untuk memudahkan proses Clustering, adapun atribut yang dirubah yaitu Nama Balita dan Jenis Kelamin, karena atribut tersebut berjenis polynominal yang nantinya tidak dapat diproses menggunakan operator k-means dengan parameter measure type Mixed Measures, Numerical Measures, dan Bregman Divergences.

Tabel 3. Parameter nominal to numerical

No	Parameter	Isi
1	Attribute Filter Type	Subset
2	Attributes	Nama Balita Jenis Kelamin
3	Coding Type	Unique Integers

Hasil atas pembacaan operator Nominal to Numerical diperoleh data yang bisa diperhatikan lewat Tabel 4 seperti:

Tabel 4 Statistika operator nominal to numerical

No	Uraian	Keterangan
1	Record	296
2	Special Attribute	1
3	Regular Attribute	5
4	Attribute	
	ID	Integer, missing 0
	Nama Balita	Nominal, missing 0
	Jenis Kelamin	Nominal, missing 0
	Umur (Bulan)	Integer, missing 0
	Berat (Kg)	Real, missing 0
	Tinggi (Cm)	Real, missing 0

4) *Data Mining*

Data mining termasuk proses supaya menemukan pola maupun data yang ada di pada data yang sudah terpilih pada saat proses transformasi, penelitian ini menggunakan *k-means* dengan mengcluster data balita, operator *k-means* bisa diperhatikan lewat Gambar 6:

Gambar 6 Operator *k-means*

Algoritma *K-means* dipakai supaya menentukan cluster dari data yang sudah dipilih, dengan melakukan terlebih dahulu pemilihan jumlah cluster yang berbeda yaitu (K = 2, 3, 4, 5, 6, 7, 8, 9, 10), jumlah *cluster* yang digunakan dalam upaya bereksperimen terhadap parameter *measure type Mixed Measures, Numerical Measures, dan Bregman Divergences*.

Tabel 5. Statistika operator *k-means*

No	Parameter	Isi
1	K	2-10
2	*MeasureTypes*	*Mixed Measures* *Numerical Measures* *Bregman Divergences*

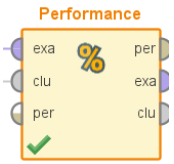
Hasil dari pembacaan operator *k-means* didapat informasi yang bisa diperhatikan lewat Tabel 6 seperti:

Tabel 6 Statistik Data *K-means*

No	Uraian	Keterangan
1	Record	296
2	Special Attribute	2
3	Regular Attribute	5
4	Attribute	
	ID	Integer, missing 0
	Cluster	Nominal, missing 0
	Nama Balita	Numeric, missing 0
	Jenis Kelamin	Numeric, missing 0
	Umur (Bulan)	Integer, missing 0
	Berat (Kg)	Real, missing 0
	Tinggi (Cm)	Real, missing 0

Penjelasan :
Special atribut yang dimaksud pada Tabel 6 adalah ID dengan tipe data *integer* dan *cluster* dengan tipe data nominal.

Menambahkan operator *Cluster Distance Performance* pada proses Clustering supaya menilai kinerja proses *Clustering* dengan metode *Davies Bouldin Index* (DBI). *Cluster Distance Performance* dipakai supaya memilih hasil dari cluster set (K = 2, 3, 4, 5, 6, 7, 8, 9, 10) dengan berdasarkan parameter *measure type Mixed Measures, Numerical Measures, dan Bregman Divergences* pada *K-means*, sampai data yang dikelompokkan mendapatkan nilai terkecil berdasarkan *Davies Bouldin Index* (DBI).

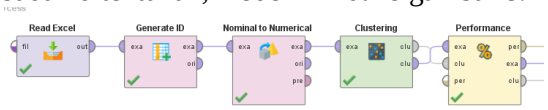


Gambar 7. Operator *Cluster Distance Performance*

Tabel 7. Parameter *Cluster Distance Performance*

No	Parameter	Isi
1	Main Criterion	Davies Bouldin Index

Data yang telah ditentukan memerlukan penggunaan satu model dalam metode *K-means*, proses model *K-means* dengan operator *Cluster Distance Performance* bertujuan untuk mendapatkan nilai *Davies Bouldin Index* (DBI) dari masing-masing cluster set (K = 2, 3, 4, 5, 6, 7, 8, 9, 10) yang sudah ditentukan, model *K-means* gambar 8.



Gambar 8 Proses Model *K-means*

Eksperimen memakai Data Mining mempermudah dan meningkatkan efektivitas proses *Clustering*. Proses ini melibatkan pengelompokkan data berdasarkan kemiripan karakteristik dari masing-masing data.

5) Evaluasi

Evaluasi termasuk langkah lanjutan dalam mencapai tujuan data mining. Proses evaluasi dibuat dengan menyeluruh yang tujuannya memastikan bahwa hasil dari tahap data mining mengikuti target yang mau diraih. Dimana saat melakukan pengujian menggunakan software RapidMiner.

Evaluasi cluster dilakukan supaya memilih total cluster yang berhasil menciptakan nilai Davies Bouldin Index terkecil, atau jumlah cluster dengan tingkat kemiripan data terdekat. Rekapitulasi Davies Bouldin Index (DBI) yang diperoleh melalui menggunakan cluster set (K = 2, 3, 4, 5, 6, 7, 8, 9, 10) parameter measure type Mixed Measures, Numerical Measures, dan Bregman Divergences bisa diperhatikan lewat Tabel 8.

Tabel 8 Hasil Evaluasi Cluster

K	Measure Type	DBI
2	Mixed Measures	0,566
	Numerical Measures	0,566
	Bregman Divergences	0,566
3	Mixed Measures	0,627
	Numerical Measures	0,626
	Bregman Divergences	0,623
4	Mixed Measures	0,692
	Numerical Measures	0,695
	Bregman Divergences	0,693
5	Mixed Measures	0,774
	Numerical Measures	0,773
	Bregman Divergences	0,767
6	Mixed Measures	0,849
	Numerical Measures	0,849
	Bregman Divergences	0,847
7	Mixed Measures	0,911
	Numerical Measures	0,890
	Bregman Divergences	0,919
8	Mixed Measures	0,899
	Numerical Measures	0,883
	Bregman Divergences	0,912
9	Mixed Measures	0,941
	Numerical Measures	0,934
	Bregman Divergences	0,969
10	Mixed Measures	0,918
	Numerical Measures	0,912
	Bregman Divergences	0,917

Mengacu pada prinsip DBI, nilai yang dianggap baik dalam data mining adalah semakin kecil atau *mendekati nol*. Oleh *Tabel 8* sebelumnya maka bisa diperhatikan jika DBI terkecil ditemukan pada K=2 dengan measure type Mixed Measures, Numerical Measures, dan Bregman Divergences dengan nilai DBI yan mendekati nol yaitu 0.566. Oleh karena itu, pengelompokkan dilakukan dengan menggunakan 2 cluster.

KESIMPULAN

Berdasarkan hasil pengelompokan data balita dengan memakai algoritma *k-means*, sehingga bisa disimpulkan bahwa Nilai K optimal berdasarkan *Davies Bouldin Index* (DBI) ditemukan pada K=2 melalui nilai DBI 0,566 yang paling mendekati. Sebab, semakin kecil nilai DBI sehingga semakin baik cluster yang diciptakan. Adapaun parameter *measure type* yang digunakan untuk menghasilkan nilai K optimal adalah *Mixed Measures, Numerical Measures*, dan *Bregman Divergences*. Adapun manfaat dalam penelitian tersebut termasuk dapat membagikan wawasan semakin dalam menyangkut situasi kesehatan serta perkembangan balita pada Desa Wanasaba Lor. Sehingga bisa menjadi landasan untuk memberikan rekomendasi dan saran kepada pihak berwenang di tingkat desa dan wilayah untuk meningkatkan kualitas hidup balita.

REFERENSI

Adiputra, I. N. M. 2021. "Clustering Penyakit Dbd Pada Rumah Sakit Dharma Kerti Menggunakan Algoritma K-Means." *INSERT: Information System and*

Emerging ….

Adiya, M. Hasmil, and Yenny Desnelita. 2019. "Jurnal Nasional Teknologi Dan Sistem Informasi Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru." 01:17–24.

Afiasari, Nur, Nana Suarna, and Nining Rahaningsi. 2023. "Implementasi Data Mining Transaksi Penjualan Menggunakan Algoritma Clustering Dengan Metode K-Means." *Jurnal SAINTEKOM* 13(1):100–110. doi: 10.33020/saintekom.v13i1.402.

Anam, Khaerul, Dadang Sudrajat, and Dian Ade Kurnia. 2022. "Analisis Segmentasi Pelanggan Menggunakan Metode K-Means Clustering." 21:273–78.

Andini, N., Taufiq, R., Priyanggodo, D. Y., & Sugiyani, Y. (2023). Penggunaan Metode Prototype pada Pengembangan Sistem Informasi Imunisasi Posyandu. *JIKA (Jurnal Informatika)*, *7*(4), 431–439. https://doi.org/10.31000/jika.v7i4.9329

Chania, Henita, Dhona Andhini, and Jaji. 2020. "Pengaruh Teknik Perkusi Dan Vibrasi Terhadap Pengeluaran Sputum Pada Balita Dengan ISPA Di Puskesmas Indralaya." *Seminar Nasional Keperawatan "Pemenuhan Kebutuhan Dasar Dalam Perawatan Paliatif Pada Era Normal Baru" Tahun 2020* 25–30.

Chusyairi, Ahmad, and Pelsri Ramadar Noor Saputra. 2019. "Pengelompokan Data Puskesmas Banyuwangi Dalam Pemberian Imunisasi Menggunakan Metode K-Means Clustering." *Telematika* 12(2):139–48. doi: 10.35671/telematika.v12i2.848.

Irawan, Yuda. 2019. "Implementation of Data Mining for Determining Majors Using K-Means Algorithm in Students of Sma Negeri 1 Pangkalan Kerinci." *Journal of Applied Engineering and Technological Science* 1(1):17–29. doi: 10.37385/jaets.v1i1.18.

Liesnaningsih, L. (2022). Rancang Bangun Sistem Informasi Tumbuh Kembang Bayi dan Balita di Posyandu Delima Kelurahan Curug Kulon. *JIKA (Jurnal Informatika)*, *6*(1), 93. https://doi.org/10.31000/jika.v6i1.5979

Kurnia, Dian Ade, and Ade Rizki Rinaldi. 2022. "Clustering Data Penjualan Produk Makanan Pada Toko Toserba Yogya Siliwangi Dengan Menggunakan Metode K-Means." 7(1):77–84.

Muhima, Rani Rotul, Muchamad Kurniawan, Septiyawan Rosetya Wardhana, Anton Yudhana, Sunardi, and Mitra Adhimukti. 2023. "An Improved Clustering Based on K-Means for Hotspots Data." *Indonesian Journal of Electrical Engineering and Computer Science* 31(2):1109–17. doi: 10.11591/ijeecs.v31.i2.pp1109-1117.

Muttaqin, Muhammad Rafi, and Meriska Defriani. 2020. "Algoritma K-Means Untuk Pengelompokan Topik Skripsi Mahasiswa." *ILKOM Jurnal Ilmiah* 12(2):121–29. doi: 10.33096/ilkom.v12i2.542.121-129.

Prakoso, Bakhtiyar Hadi, Ervina Rachmawati, Demiawan Rachmatta Putro Mudiono, Veronika Vestine, and Gandu Eko Julianto Suyoso. 2023. "Klasterisasi Puskesmas Dengan K-Means Berdasarkan Data Kualitas Kesehatan Keluarga Dan Gizi Masyarakat." *Jurnal Buana Informatika* 14(01):60–68. doi: 10.24002/jbi.v14i01.7105.

Rahmat, C. A., H. Permatasari, and ... 2023. "Penerapan K-Means Untuk Clustering Kondisi Gizi Balita Pada Posyandu." *Jurnal Media …*.

Saputro, D. T., and W. P. Sucihermayanti. 2021. "Penerapan Klasterisasi Menggunakan K-Means Untuk Menentukan Tingkat Kesehatan Bayi Dan Balita Di Kabupaten Bengkulu Utara." *Jurnal Buana Informatika*.

Tinendung, Intan Saleha, and Ilka Zufria. 2023. "Pengelompokan Status Stunting Pada Anak Menggunakan Metode K-Means Clustering." 7:2014–23. doi: 10.30865/mib.v7i4.6908.