



CLUSTERING STUNTING PADA BALITA DENGAN METODE K-MEANS DI PUSKESMAS KANATANG

Rafael Ranjawali¹, Alfrian Carmen Talakua², Reynaldi Thimotius Abineno³

Program Studi Teknik Informatika, Fakultas Sains Dan Teknologi, Universitas Kristen Wira Wacana Sumba

Jl.R.Suprato No.35,Prailiu,Kec.Kota Waingapu,Kabupaten Sumba Timur,Nusa Tenggara Timu

Corresponding author: alfriantalakua@unkriswina.ac.id¹, rafaelranjawali2@gmail.com²
reynaldi@unkriswina.ac.id³

ABSTRACT

Stunting is a chronic nutritional condition in toddlers associated with body growth rate, measured by length or height, and can occur since the fetus is still in the womb. Kanatang Health Center is one of the health centers contributing to stunting cases in Indonesia, which is located in East Sumba Regency with quite high stunting cases. Based on data from August 2022, it was at 9.2% of stunting cases, increasing in February 2023, reaching 9.5%. So, an approach is needed to classify data on toddlers who experience stunting based on the factors that cause it using the K-Means method. The purpose of the study is to provide support to the government in taking the right policy to reduce the incidence of stunting in toddlers. Based on the results of testing the Davies Bouldin Index value of calculations using the RapidMiner application, the optimal performance is $K = 2$ with a value of 0.537 where the value is close to 0, it can be said that the evaluation results produce good clusters, cluster 0 is a cluster with a very high stunting causal factor of 54 toddlers, and Cluster 1 is a group consisting of 42 toddlers with a low level of stunting causal factors.

Keywords: Stunting, K-Means, RapidMiner, Clustering, Toddler.

ABSTRAK

Stunting adalah kondisi gizi kronis pada balita yang terkait dengan tingkat pertumbuhan tubuh, yang diukur berdasarkan panjang atau tinggi badan, dan bisa terjadi sejak janin masih dalam kandungan. Puskesmas Kanatang merupakan salah satu puskesmas penyumbang kasus stunting di Indonesia, yang berada di Kabupaten Sumba Timur dengan kasus stunting yang cukup tinggi. Berdasarkan data bulan Agustus 2022 berada pada angka 9,2% kasus stunting, meningkat pada bulan Februari 2023, mencapai 9,5%. Sehingga, dibutuhkan suatu pendekatan untuk mengklasifikasikan data balita yang mengalami stunting berdasarkan faktor-faktor yang menyebabkannya dengan menggunakan metode K-Means. Tujuan penelitian untuk memberikan dukungan kepada pemerintah dalam mengambil kebijakan yang tepat untuk mengurangi angka kejadian stunting pada balita. Berdasarkan hasil pengujian nilai Davies Bouldin Index perhitungan menggunakan aplikasi RapidMiner, menghasilkan performa yang optimal yaitu $K=2$ dengan nilai 0.537 dimana nilai tersebut mendekati 0 maka dapat dikatakan hasil evaluasi menghasilkan klaster yang baik, cluster 0 merupakan cluster dengan faktor penyebab stunting sangat tinggi yaitu 54 balita, dan Cluster 1 adalah kelompok yang terdiri dari 42 balita dengan tingkat faktor penyebab stunting yang rendah.

Kata kunci: Stunting, K-Means, RapidMiner, Clustering, Balita.

PENDAHULUAN

Pertumbuhan fisik pada balita adalah hasil dari faktor-faktor yang akumulasi sejak lahir. Kondisi gizi yang mencukupi dan kesehatan yang baik selama masa balita memiliki peran sangat penting dalam menentukan kesehatannya di masa mendatang. Namun, keadaan

di mana zat gizi, terutama protein dan energi, tidak terpenuhi dengan baik pada balita dapat menyebabkan gangguan dalam pertumbuhan dan perkembangannya (Yulia et al., 2021).

Kurang gizi dan *stunting* pada balita adalah dua permasalahan yang saling berkaitan. Hal ini bisa menimbulkan gangguan perkembangan fisik, mempengaruhi kemampuan mental, dan belajar tidak maksimal, serta prestasi belajar yang buruk. Selain itu, efek jangka panjang yang diakibatkan dari *stunting* dan penyebab lain seperti kurang gizi, sering kali dianggap sebagai salah satu faktor risiko diabetes, hipertensi, obesitas dan kematian akibat infeksi (Faujia et al., 2022).

Di Indonesia, masalah *stunting* pada balita dianggap sebagai isu kesehatan masyarakat yang serius, dan berada dalam kategori tingkat prevalensi yang sangat tinggi menurut standar WHO. Hal ini menunjukkan bahwa frekuensinya adalah 37,2%, naik dari 35,6% pada tahun 2019 dan 36,8% pada tahun 2020, menurut Kementerian Kesehatan Republik Indonesia. Provinsi penyumbang kasus *Stunting* tertinggi di Indonesia adalah Provinsi Nusa Tenggara Timur (NTT), Salah satu kabupaten di NTT yang memiliki tingkat *stunting* yang cukup tinggi adalah Kabupaten Sumba Timur, dengan prevalensi *stunting* sekitar 14,9% pada bulan agustus tahun 2022, berdasarkan data dari Kemenkes. Data tersebut berdasarkan laporan dari masing-masing Puskesmas sebagai hasil rekapan pada bulan Agustus 2022.

Puskesmas Kanatang merupakan puskesmas yang berada di Kabupaten Sumba Timur. Salah satu Puskesmas yang memiliki kasus *stunting* yang sangat tinggi dari puskesmas lainnya, pada tahun 2023 naik 0,3%, berdasarkan data bulan Agustus 2022 berada pada angka 9,2% kasus *stunting* dan mengalami peningkatan pada bulan Februari 2023 dengan pencapaian 9,5% dengan jumlah kasus *stunting* sebanyak 96 balita. Pola pertumbuhan *stunting* pada balita dipengaruhi oleh faktor-faktor seperti, Asupan nutrisi yang tidak memadai, terutama zat-zat penting seperti protein, vitamin, dan mineral, dapat menyebabkan pertumbuhan terhambat pada balita, Pola makan yang tidak seimbang dan tidak mencukupi dapat menyebabkan kurangnya asupan nutrisi yang diperlukan untuk pertumbuhan yang optimal, Kesehatan ibu selama kehamilan dan periode menyusui dapat memengaruhi perkembangan janin dan kualitas ASI (Air Susu Ibu), Lingkungan yang tidak bersih dan kurangnya akses terhadap air bersih serta fasilitas sanitasi yang layak dapat meningkatkan risiko infeksi dan gangguan absorpsi nutrisi, Pelayanan kesehatan yang tidak memadai, termasuk pemantauan pertumbuhan, pendampingan nutrisi, dan edukasi kepada orang tua, dapat berdampak pada pola pertumbuhan *stunting*, Kondisi sosial dan ekonomi keluarga juga dapat memengaruhi akses terhadap nutrisi yang baik dan layanan kesehatan.

Dari persoalan diatas, pemerintah mempunyai peran yang sangat penting dalam menekan peningkatan kasus *stunting* di Puskesmas Kanatang dengan mengambil kebijakan yang tepat bahwa kasus ini menjadi fokus utama untuk kemajuan bangsa kedepannya. Sehingga perlu melakukan peningkatan pelayanan pemeriksaan di puskesmas merupakan salah satu yang dapat dilakukan, atau dengan menambah bantuan berupa vitamin kepada balita dan ibu hamil.

Agar kebijakan yang diambil oleh pemerintah dapat tepat sasaran, maka perlu dilakukan pengklasteran pada balita yang tergolong *stunting* dengan penentuan dari beberapa faktor yang saling berkaitan yaitu data usia balita, status gizi, penyebab *stunting*. Pengklasteran perlu dilakukan berdasarkan beberapa indikator seperti status gizi pada anak, yang menunjang dalam upaya penurunan tingkat kasus *stunting*. Salah satu alat statistik yang dapat digunakan untuk melakukan pengklasteran terhadap kasus *stunting* di Puskesmas Kanatang adalah analisis kluster.

MATERI DAN METODE

Data mining

Data mining adalah proses penemuan pola yang relevan dan diinginkan dalam basis data yang besar, dengan tujuan untuk membantu pengambilan keputusan di masa mendatang. Pola-pola ini dapat diidentifikasi oleh algoritma khusus yang mampu menyajikan analisis data yang bernilai, dan kemudian dapat dipelajari secara mendalam, sering kali dengan menggunakan alat bantu pengambilan keputusan lainnya. (Sari, 2021).

Kata *Data Mining* diklaim pula *Knowledge Discovery* atau penggalian pengetahuan dan pengenalan pola adalah dua istilah yang memiliki akurasi dan kegunaan yang berbeda-beda. Istilah "*Knowledge Discovery*" atau penggalian pengetahuan digunakan dengan tepat karena tujuan utamanya dari *data mining* adalah untuk mengungkapkan pengetahuan yang tersembunyi dari kumpulan data. Sedangkan istilah "*Pattern Recognition*" atau pengenalan pola tepat untuk digunakan karena pengetahuan yang ingin ditemukan memang berbentuk pola-pola yang mungkin perlu diidentifikasi lebih lanjut dari kumpulan data yang sedang dihadapi (Nainggolan dan Purba, 2019).

Data Mining sangat menarik perhatian industri informasi dalam beberapa tahun terakhir karena adanya kelimpahan data dan kebutuhan yang semakin meningkat untuk mengolah data tersebut menjadi informasi dan pengetahuan yang berharga. *Data Mining* juga merupakan gabungan dari berbagai disiplin ilmu, termasuk Basis Data, Kecerdasan Buatan, Statistik, dan pembelajaran mesin, serta lainnya (Yulia, 2021).

Clustering

Manurut Anggraeni *et al.*, (2023), *clustering* adalah metode pengelompokan data yang digunakan untuk mengenali kelompok-kelompok (*cluster*) yang dihasilkan dari pengelompokan unsur-unsur yang lebih kecil berdasarkan adanya kemiripan satu sama lain. Kemiripan yang menjadi dasar pengelompokan tidak bersifat universal sehingga ukuran-ukuran penyamanya harus dijabarkan terlebih dahulu oleh peneliti atau penganalisis. *Clustering* merupakan metode pengelompokan data yang sering digunakan sebagai satu di antara metode *data mining* atau penggalian data (Faujia *et al.*, 2022).

Clustering adalah proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan *cluster*. Maka itu, metode *clustering* ini berguna untuk menemukan kelompok yang tidak dikenal dalam data. Dalam *business intelligence*, *clustering* bisa mengelompokkan banyak *customer* ke beberapa kelompok. Contohnya mengelompokkan *customer* ke dalam beberapa *cluster* dengan kesamaan karakteristik yang kuat. *Clustering* ini juga dikenal sebagai data segmentasi karena *clustering* mempartisi banyak dataset ke dalam banyak grup berdasarkan kesamaannya (Fadilah *et al.*, 2022).

Stunting

Stunting merupakan kondisi gagal pertumbuhan pada balita (anak di bawah lima tahun) yang disebabkan oleh kekurangan gizi kronis, sehingga menyebabkan balita memiliki tinggi badan yang terlalu pendek untuk usianya. Kekurangan gizi ini dapat terjadi sejak balita berada dalam kandungan dan pada masa awal setelah bayi lahir, namun gejala *stunting* baru menjadi terlihat setelah balita mencapai usia 2 tahun (Faujia *et al.*, 2022). *Stunting* merupakan isu kesehatan masyarakat yang terkait dengan peningkatan risiko terkena penyakit, kematian, serta gangguan pertumbuhan baik fisik maupun kognitif.

Indikator Tinggi Badan Ideal

Tabel 3 menyajikan indikator tinggi badan ideal anak berdasarkan standar Kementerian Kesehatan RI dan WHO (World Health Organization).

Table 1. Standar Anthropometri

No	Standar Anthropometri (Tinggi Badan Ideal Berdasarkan Usia)
1.	Bayi Usia 0 - 3 Bulan: Tinggi Badan 40,4 – 60 cm
2.	Bayi Usia 4 - 6 Bulan: Tinggi Badan 60,5 – 66,0 cm
3.	Bayi Usia 7 – 9 Bulan: Tinggi Badan: 67,5 – 70,5 cm Bayi Usia 10 – 12 Bulan: Tinggi Badan: 72 – 74,5 cm
4.	cm
5.	Balita Usia 13 – 24 Bulan: Tinggi Badan: 82 – 92 cm
6.	Balita Usia 25 – 32 Bulan: Tinggi Badan: 83 – 95 cm
7.	Balita Usia 33 – 44 Bulan: Tinggi Badan: 84 – 97 cm
8.	Balita Usia 44 – 58 Bulan: Tinggi Badan: 85 – 98 cm

Metode *K-Means*

K-Means adalah salah satu algoritma yang termasuk dalam kategori unsupervised learning. Algoritma *K-Means* berfungsi untuk mengelompokkan data ke dalam berbagai cluster. Keunggulan dari algoritma ini adalah kemampuannya untuk menerima data tanpa adanya label kategori sebelumnya. Dengan menggunakan metode ini, data dapat dikelompokkan berdasarkan kesamaan fitur atau atribut tertentu tanpa memerlukan bantuan label atau pengawasan sebelumnya. Matdoan *et al.*, (2022), menyatakan bahwa *K-Means Clustering* Algoritma juga merupakan metode *non-hierarchy*. Metode *Clustering* Algoritma adalah mengelompokkan beberapa data ke dalam kelompok yang menjelaskan data dalam satu kelompok memiliki karakteristik yang sama dan memiliki karakteristik yang berbeda dengan data yang ada di kelompok lain. *Cluster Sampling* adalah teknik pengambilan sampel di mana unit-unit populasi dipilih secara acak dari kelompok yang sudah ada yang disebut *cluster*, jadi *Clustering* atau klasterisasi adalah salah satu masalah yang menggunakan teknik *unsupervised learning* (Yulia, 2021).

Davies-Bouldin Index

Abineno, (2022) menyatakan Indeks Davies-Bouldin (*Davies-Bouldin Index*) adalah metode evaluasi yang digunakan untuk mengukur kualitas setiap kluster dalam pengelompokan data. Metode ini mengevaluasi seberapa baik setiap objek dikelompokkan ke dalam suatu kluster. *Davies-Bouldin Index* menggabungkan dua konsep, yaitu *cohesion* (kekompakan) dan *separation* (pemisahan). *Cohesion* mengukur sejauh mana setiap data berada dekat dengan pusat kluster tempat data tersebut berada, sedangkan *separation* mengukur jarak antara pusat kluster (*centroid*) dari kluster yang berbeda. Metode ini digunakan untuk mengevaluasi performa dan mengukur akurasi dalam menggunakan metode *K-Means*.

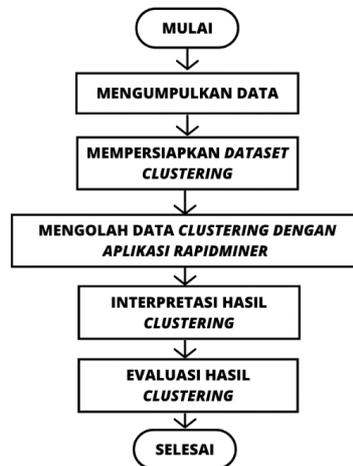
RapidMiner

RapidMiner adalah perangkat lunak analisis data yang digunakan untuk mengambil informasi dari data. *RapidMiner* dapat digunakan untuk mengumpulkan data dari berbagai sumber, termasuk *database*, *file Excel*, dan data streaming. Setelah data dikumpulkan, *RapidMiner* dapat digunakan untuk menganalisis data dengan menggunakan berbagai teknik analisis, termasuk *clustering*, klasifikasi, regresi, dan lain sebagainya (Wibowo et al., 2022).

Data yang digunakan dalam penelitian dicatat dalam aplikasi bernama e-PPGBM. Data tersebut yang dapat dilihat Data yang dikumpulkan dapat berupa data usia balita, tinggi

badan, berat badan, nama balita, tanggal lahir, serta faktor-faktor lain yang berpotensi mempengaruhi kondisi *stunting* pada balita di Puskesmas Kanatang.

Berikut merupakan kerangka berpikir yang akan menggambarkan tahapan-tahapan yang akan dilakukan.



Gambar 1. Tahapan Penelitian

Tahap pertama perlu dilakukan metode pengumpulan data yang digunakan dalam penelitian ini adalah dengan menggunakan metode penelitian kualitatif. Penelitian akan berfokus pada hasil-hasil wawancara dan observasi, laporan, jurnal, berita maupun hasil penelitian serupa yang telah dihasilkan pada sebelumnya. Pengumpulan data dilakukan dengan cara melakukan survei atau mengunjungi secara langsung di lokasi objek penelitian yaitu di Puskesmas Kanatang. Terdapat jenis data yang diperoleh yaitu berupa format data excel. Penelitian ini menggunakan data pada bulan Februari 2023 sebanyak 96 balita yang diperoleh dari Puskesmas Kanatang. Data yang digunakan dalam penelitian dicatat dalam aplikasi bernama e-PPGBM.

Pada tahap ini diperlukan mempersiapkan *dataset*, karena setelah melakukan pengumpulan data, perlu dilakukan pengecekan kembali data-data yang sudah terkumpul. Kemudian melakukan pengolahan data atau melakukan analisis kluster lebih cepat karena sudah meminimalisir terjadinya kesalahan data. *Dataset* dapat terdiri dari berbagai jenis data, seperti data numerik (angka), data kategorikal (kelas), atau data teks, serta membersihkan tabel-tabel kosong lainnya. Setelah *dataset* sudah disiapkan maka akan lebih mempermudah dalam pengolahan atau *cluster*.

Pada tahap ini, *clustering* dilakukan dengan menggunakan metode *clustering* yang telah dipilih yaitu *K-Means*. Perhitungan menggunakan aplikasi *RapidMiner* untuk analisis data *stunting*, dari *Dataset* yang sudah disiapkan.

Hasil *clustering* yang diperoleh kemudian diinterpretasikan dengan cara menganalisis karakteristik atau ciri-ciri yang dimiliki oleh setiap kelompok data yang terbentuk, hasil tersebut disajikan dalam bentuk Tabel, dan Grafik. Hal ini dapat membantu dalam mengidentifikasi faktor-faktor yang berpotensi mempengaruhi kondisi *stunting* pada balita.

Metode evaluasi yang digunakan untuk mengukur kualitas setiap kluster dalam pengelompokan data adalah *Davies-Bouldin Index*. Metode ini digunakan untuk mengevaluasi performa dan mengukur akurasi dalam menggunakan metode *K-Means*.

HASIL DAN PEMBAHASAN

Pengumpulan Data

Pengumpulan data telah dilakukan Di Puskesmas Kanatang terkait data stunting pada balita, terdapat jenis data yang diperoleh yaitu berupa format data excel yang didapat dari Puskesmas Kanatang. Penelitian ini menggunakan data pada bulan Februari 2023 sebanyak 96 balita yang diperoleh dari Puskesmas Kanatang. Pengolahan data akan dilakukan dengan dataset yang sudah disiapkan. Data yang digunakan dalam penelitian dicatat dalam aplikasi Bernama e-PPGBM.

Mempersiapkan Dataset

Mempersiapkan *dataset* sangat penting, karena setelah melakukan pengumpulan data, perlu dilakukan pengecekan kembali data-data yang sudah terkumpul. Berikut *dataset* yang sudah disiapkan untuk pengolahan *clustering* menggunakan aplikasi *Rapidminer*.

Tabel 1. *Dataset*

No	Nama Balita	Desa	Usia Saat Ukur	Berat	Tinggi
1.	Ambrafael U Lingu	TEMU	16	7,9	76,4
2.	Silvester Stiven Modu	KUTA	40	12	90
3.	Evelin M. Bayo	KUTA	14	9,1	70,6
4.	Silfester Alfero	KUTA	38	8,7	75,8
5.	Bela Vanaia K. W	KUTA	56	13,5	94
6.	Erlando L. Leijo	TEMU	20	7,6	74,5
....
95	Into Yowa Jurumana	TEMU	23	9,6	77,2
96.	Aritwan	TEMU	23	9,0	70,0

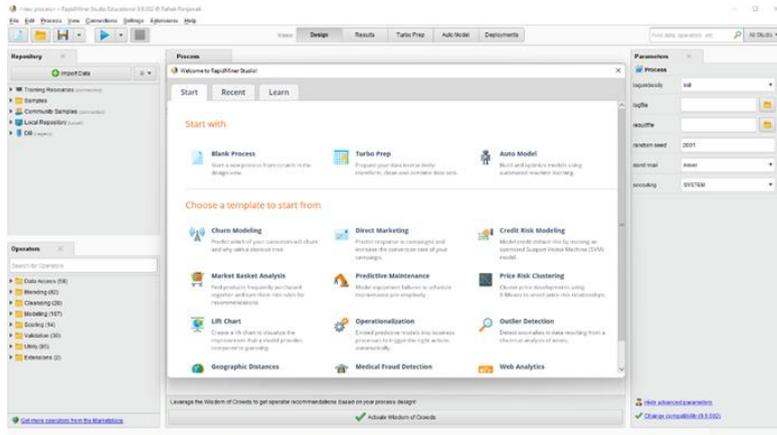
Berdasarkan dataset yang didapatkan dari hasil pengumpulan data, maka data yang digunakan adalah data dengan atribut nama balita, desa, usia saat diukur, BB/U, dan TB/U. Nama balita merupakan balita yang dikategorikan stunting, Desa merupakan atribut yang berisi 6 desa, usia saat diukur berisikan usia balita stunting, sedangkan BB/U dan TB/U berisi status gizi berdasarkan nilai *antropometri* BB/U dan TB/U.

Clustering Menggunakan Perangkat Lunak *RapidMiner*.

Data yang diperoleh digunakan sebagai input untuk membangun model algoritma K-Means Clustering menggunakan perangkat lunak *RapidMiner*.

Tampilan Menu Utama *RapidMiner*

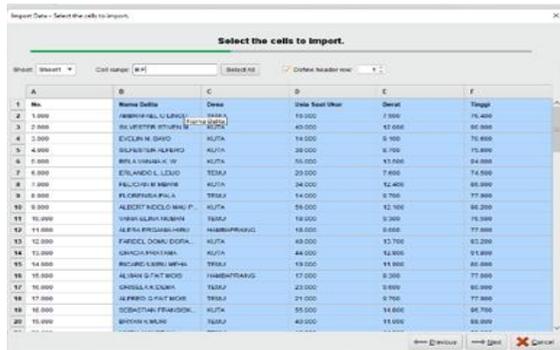
Menu utama *RapidMiner* memiliki tampilan awal yang mencakup berbagai komponen dan fitur. Salah satunya adalah "*New Process*", yang merupakan langkah pertama dalam menggunakan *RapidMiner*.



Gambar 2. Tampilan Menu Utama *RapidMiner* Versi 9.9

Tampilan *Select The Cells To Import*

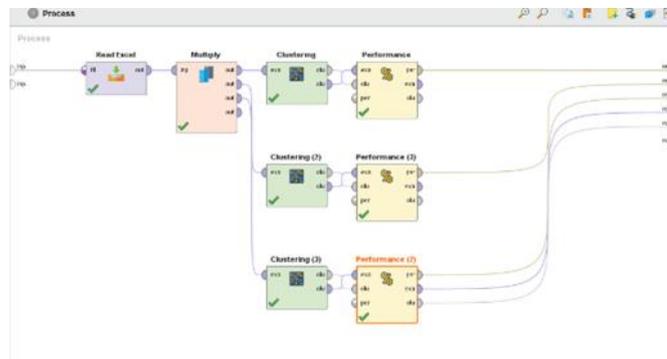
Sistem memberikan instruksi tentang bagaimana cara mengimpor data baru yang akan diolah selanjutnya, dengan format data saat ini dalam bentuk *Excel*. Tahapan ini terlihat dalam gambar berikut ini.



Gambar 3. Tampilan *Select The Cells To Import*

Tampilan *Clustering RapidMiner*

Algoritma *K-Means* akan diterapkan dalam proses yang telah dibentuk untuk melakukan klusterisasi atau pengelompokan. Operator klusterisasi yang digunakan adalah *K-Means*. Setelah pembuatan model selesai, proses akan dijalankan untuk mendapatkan hasil klusterisasi. Gambar 4. menunjukkan model proses terbaik yang menggunakan algoritma *K-Means*.



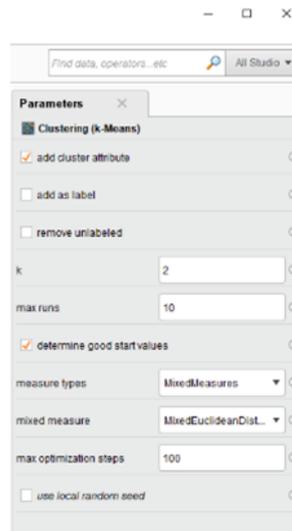
Gambar 4. Tampilan *Clustering RapidMiner*

Dari gambar diatas, menunjukkan beberapa operator seperti *Read Excel* dimana *dataset* akan dimasukkan dalam *Rad Excel*, *Multiply* yang berfungsi untuk menjalankan satu

atau lebih operator, operator *cluster* yaitu algoritma *k-means*, kemudian *performance* untuk evaluasi kinerja dari *k-means*.

Penentuan jumlah kluster

Langkah selanjutnya adalah melakukan pengaturan *algoritma K-Means* melalui menu "*Parameter Clustering K-Means*", berikut tampilan penentuam jumlah parameter kluster.



Gambar 5. Parameter Jumlah Kluster

Berdasarkan Gambar 5, perlu memasukkan nilai *K* yang diinginkan. Dalam iterasi yang dilakukan sebanyak 10 kali, nilai *K* yang dimasukkan adalah $k=2,3,4$. Tujuan iterasi tersebut adalah untuk menemukan kluster dengan nilai *DBI* yang mendekati 0 atau yang paling kurang, maka dianggap sebagai kluster terbaik. Setelah iterasi selesai, Proses klusterisasi telah dilakukan untuk mendapatkan hasil dari algoritma *K-Means* yang diimplementasikan melalui aplikasi *RapidMiner*. Setelah melalui beberapa iterasi, ditemukan bahwa nilai optimal untuk jumlah kluster (*K*) adalah 2. Hasil klusterisasi akhir diperoleh menggunakan algoritma *K-Means* berdasarkan nilai *K* yang telah ditentukan tersebut.

Interpretasi Hasil

Analisis kluster *K-Means* dilakukan untuk memisahkan balita ke dalam kelompok-kelompok berdasarkan variabel-variabel terkait gizi seperti usia balita berat badan, tinggi badan, dan lain-lain. Analisis kluster dapat membantu mengidentifikasi pola pola tertentu yang mengindikasikan gizi buruk atau masalah kesehatan lainnya. Misalnya, kluster tertentu memiliki tinggi badan lebih rendah dan berat badan kurang, menunjukkan adanya masalah stunting dan *underweight* pada kelompok tersebut. Dengan mengelompokkan balita berdasarkan status gizi, dapat membantu mengidentifikasi faktor-faktor penyebab yang berkontribusi terhadap masalah gizi. Misalnya, kelompok dengan status gizi buruk mungkin berbagi faktor risiko tertentu seperti pola makan yang tidak seimbang atau akses terbatas terhadap nutrisi. Setelah dilakukan proses clustering, data tersebut telah dikelompokkan menjadi dua cluster yang diberi nama Cluster 0 dan Cluster 1. Hasil kluster pada setiap balita akan ditampilkan dalam bentuk Tabel 2.

Tabel 2. Anggota Hasil Kluster

Cluste r	Berdasarkan Usia Balita	Jumlah
0	31 Bulan - 57 Bulan	42

1	0 Bulan – 30 Bulan	54
	TOTAL	96

Tabel 2 menunjukkan hasil perhitungan menggunakan aplikasi RapidMiner, yang telah dibagi menjadi 2 cluster. Cluster 1 terdiri dari 54 balita dengan faktor penyebab stunting tinggi, dan usia balita dalam cluster ini berkisar antara 0 bulan hingga 30 bulan. Sedangkan, cluster 0 terdiri dari 42 balita dengan faktor penyebab stunting rendah, dan usia balita dalam cluster ini berkisar antara 31 bulan hingga 57 bulan. Perlu dicatat bahwa kedua kelompok balita ini diklasifikasikan berdasarkan Standar Antropometri, yaitu standar ideal untuk balita.

Aggregat Value Clustering

Setelah melakukan proses pengelompokan data (*clustering*) berdasarkan tabel di atas, selanjutnya akan ditentukan nilai minimum (*min*), maksimum (*max*), dan rata-rata (*mean*) dari setiap cluster dan variabel. Tujuan dari langkah ini adalah untuk mengidentifikasi karakteristik dari masing-masing cluster. Berikut adalah hasil dari proses perhitungan *Aggregate Value*, dapat dilihat pada Tabe 3:

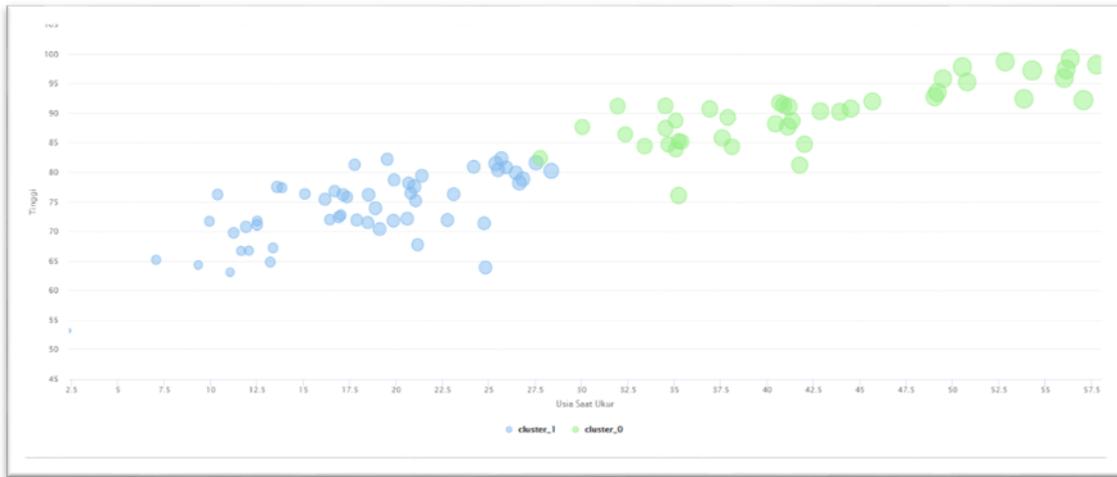
Tabel 3. Aggregate Value Clustering

MAX			
Cluster	Usia Saat Ukur	Berat Badan	Tinggi Badan
0	57,00	15,00	98,20
1	30,00	11,90	83,30
MIN			
Cluster	Usia Saat Ukur	Berat Badan	Tinggi Badan
0	31,00	8,00	75,80
1	3,00	6,00	51,20
MEAN			
Cluster	Usia Saat Ukur	Berat Badan	Tinggi Badan
0	27,24	11,98	89,58
1	18,22	8,70	73,86

Setiap *cluster* memiliki karakteristik yang berbeda berdasarkan variabel Usia Saat Ukur Badan, Berat Badan, dan Tinggi Badan. Berdasarkan hasil *Aggregate Value* diatas, nilai tertinggi (*Max*) yang ditemukan pada *Cluster 0*, terdapat pada variabel Usia Saat Ukur dengan nilai maksimal 57,00%, untuk variabel Berat Badan adalah 15,00% dan 98,20% untuk nilai maksimal pada variabel Tinggi Badan. Nilai terendah (*Min*) yang terdapat pada *Cluster 1*, yaitu dalam variabel Usia Saat Ukur adalah 3,00%, untuk variabel Berat Badan dengan nilai 6,00%, sedangkan variabel Tinggi Badan adalah 51,20%. Nilai rata-rata (*Mean*) dari *Cluster 0*, pada variabel Usia Saat Ukur adalah 27.24%, sedangkan nilai mean yang paling rendah dari variabel Berat Badan dengan nilai 8.70%, dan 89.58% untuk nilai rata-rata dari variabel Tinggi Badan.

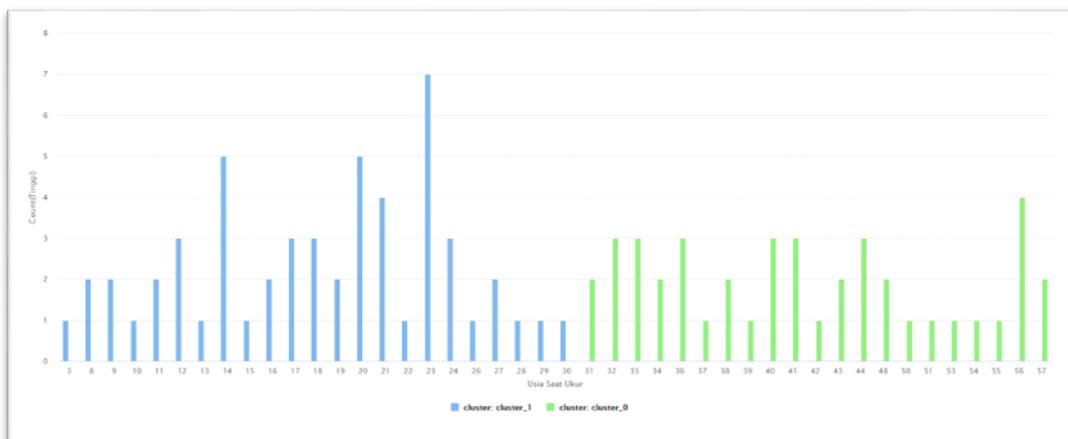
Visualisasi Hasil Clustering

Hasil pengujian data menggunakan perangkat lunak *RapidMiner 9.9* menghasilkan beberapa *output*. Dalam gambar visualisasi, terdapat garis yang menunjukkan jumlah kelompok klaster untuk kasus stunting di Puskesmas Kanatang. Visualisasi tersebut menampilkan sumbu x, sumbu y, dan warna kustom (klaster) di sebelah bawah, yang semuanya terlihat dalam gambar di bawah ini.



Gambar 6. *Scatter Plot* Usia(x), Tinggi Badan (y)

Gambar 6. adalah gambaran *scatter plot* dari Usia saat diukur yang sebagai variabel x dan Tinggi Badan sebagai variabel y. sehingga dapat dilihat bahwa kasus *stunting* paling tinggi berada pada cluster 1 yaitu berdasarkan usia balita. Begitu juga pada Gambar 7. *Bar Plot*.

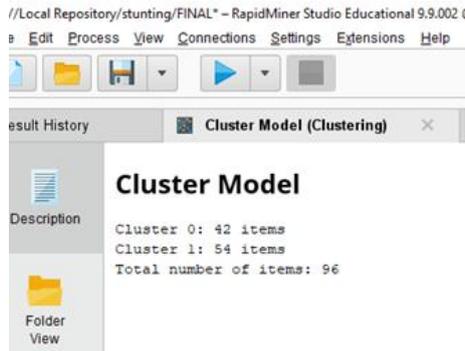


Gambar 7. *Bar Plot* Usia(x), Tinggi Badan (y)

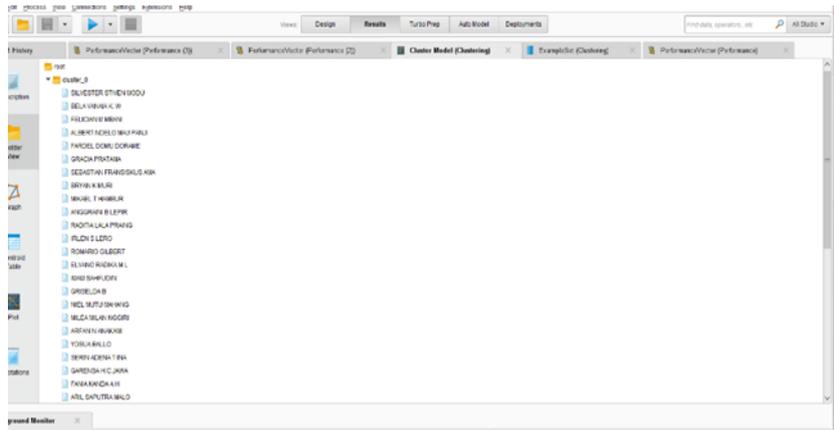
Gambar 6. dan Gambar 7. Menunjukkan hasil akhir dari proses klustering data menunggunakan aplikasi *Rapidminer* yang digunakan untuk melakukan pengelompokan data klaster. Berdasarkan hasil pengujian, dapat disimpulkan bahwa terdapat 42 balita dalam klaster 0 yang ditandai dengan warna hijau, sementara klaster 1 memiliki 54 balita yang ditandai dengan warna biru. Hasil dari clustering kemudian divisualisasikan menggunakan *Bar Plot* dan *Scatter Plot*, yang serupa dengan contoh yang terlihat pada gambar.

Cluster Model

Dalam model klaster (*clustering*), terdapat beberapa tampilan yang menampilkan hasil klaster, salah satunya adalah tampilan teks (*text view*) yang menunjukkan hasil pengelompokan berdasarkan klaster dan jumlah anggota dalam setiap klaster. Klaster 0 memiliki 42 item, sedangkan klaster 1 memiliki 54 item. Tampilan teks tersebut dapat dilihat dalam Gambar 8.



Gambar 8. Cluster Model



Gambar 9. Folder View

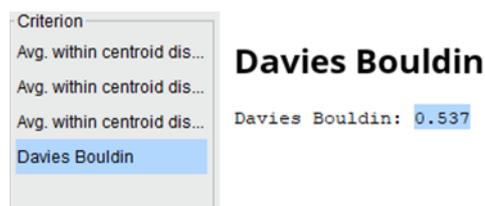
Dari Gambar 9, *Folder view* menunjukkan data secara komprehensif dari bagian-bagian kluster, dengan setiap anggota dari dua kluster menampilkan nama balita. Setelah menerapkan pengelompokan kasus *stunting* menggunakan *RapidMiner*.

Evaluasi Hasil Clustering

Berdasarkan Tabel 2. dibawah ini, dengan menggunakan metode k-means clustering dan didukung oleh perhitungan Davies-Bouldin Index (DBI), dilakukan 3 iterasi percobaan untuk menentukan cluster yang paling optimal berdasarkan nilai DBI terbaik.

Tabel 4. Nilai *Davies-Bouldin Index (DBI)*

K	Avg. within centroid distance	DBI
2	82.117	0.537
3	47.028	0.573
4	30.133	0.630



Gambar 10. Nilai *DBI* Iterasi Ke 2

Gambar 10. memberikan informasi tentang nilai *DBI* dan *performa* klusterisasi. Informasi ini diperoleh melalui pengujian dengan menggunakan perhitungan *Davies-Bouldin Index (DBI)* menggunakan *RapidMiner*. Hasil terbaik dari *DBI* ditemukan ketika $K=2$, dengan

nilai DBI sebesar 0.537 yang mendekati 0. Ini menunjukkan bahwa Semakin kecil nilai Davies Bouldin Index (DBI) yang dihasilkan, semakin baik hasil klasterisasi yang terbentuk. DBI digunakan untuk mengukur kualitas dari pengelompokan data, dan nilai yang lebih kecil menandakan bahwa pembagian data ke dalam kelompok-kelompok tersebut lebih baik dan lebih terpisah satu sama lain. Dengan demikian, DBI adalah indikator penting dalam mengevaluasi keberhasilan metode klasterisasi K-Means dalam proses pengelompokan data.

KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, hasil pengujian data menggunakan *tools RapidMiner* pada pengelompokan balita Stunting di Puskesmas Kanatang Per akhir bulan februari 2023 menggunakan metode *K-Means Clustering* dengan tiga variabel yaitu usia, berat badan (BB), dan tinggi badan (TB). Sehingga menghasilkan 2 *cluster*, dimana *cluster* penyebab *stunting* tinggi merupakan *cluster* 1 yang terdiri dari 54 balita, dan *Cluster* 0 adalah kelompok dengan tingkat *stunting* yang rendah, terdiri dari 42 balita. Berdasarkan Standar *Antropometri* standar ideal balita. Sehingga perlu adanya pendampingan dari posyandu serta puskesmas terkait *stunting* kepada orang tua balita sehingga balita yang termasuk *stunting* bisa berkurang bahkan tidak ada lagi di bulan berikutnya. Berdasarkan hasil pengujian menggunakan Davies Bouldin Index (DBI), perhitungan dilakukan dengan *Tools RapidMiner* sehingga menghasilkan nilai evaluasi *performa* yang baik, adalah $K=2$ memiliki nilai 0.537 dimana nilai tersebut mendekati 0 yang berarti klaster yang di evaluasi dengan hasil klaster yang bagus. Dari hasil analisis klastering ini, Pemerintah diharapkan dapat meningkatkan akses terhadap pelayanan kesehatan berkualitas, termasuk pemantauan pertumbuhan dan pendampingan nutrisi bagi ibu hamil dan balita. Serta memberikan edukasi kepada orang tua tentang pentingnya nutrisi yang seimbang dan praktik makan yang baik, pemberian suplemen gizi, dan Promosi praktik pemberian ASI eksklusif dan menyusui yang baik.

SARAN

Berdasarkan hasil penelitian, perlu di lakukan pengembangan lebih lanjut diperlukan penggabungan algoritma K-Means dengan algoritma lain. Selain itu, disarankan untuk meningkatkan jumlah data yang digunakan dengan menambahkan lebih banyak data. Penting juga untuk mempertimbangkan pendekatan yang lebih rinci dalam penelitian berikutnya. Hal ini dapat dilakukan dengan menggunakan data yang dibagi berdasarkan kecamatan di Kabupaten atau Kota, serta mempertimbangkan kasus Stunting di Puskesmas Kanatang dalam bulan-bulan berikutnya. Terdapat usia-usia tertentu yang perlu di perhatikan dan memberikan penanganan khusus, Salah satu pendekatan yang bisa dipertimbangkan adalah menggunakan algoritma berbeda, seperti K-Medoids atau Fuzzy C-Means. Dalam upaya memvalidasi hasil, penggunaan metode lain seperti Silhouette atau metode lainnya juga bisa diimplementasikan. Lebih lanjut, untuk mengenali perbedaan hasil evaluasi DBI (Davies-Bouldin Index), penelitian selanjutnya dapat menggali informasi tentang perbedaan antara evaluasi manual dan penggunaan tools seperti RapidMiner dalam menghitung indeks tersebut.

DAFTAR PUSTAKA

- Abineno, R. T. (2022). Clustering Dampak Dan Penanganan Covid 19 Se-Asia Menggunakan Metode K-Means Dengan Variabel-Variabel Pada Epidemiologi (Thesis, Universitas Atma Jaya Yogyakarta).
- Anggraeni, M. R., Yudatama, U., & Maimunah, M. (2023). Clustering Prevalensi Stunting Balita Menggunakan Agglomerative Hierarchical Clustering. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 7(1), 351-359.
- Fadilah, A., Pangestu, M. N., Lumbanbatu, S., & Defiyanti, S. (2022). Pengelompokan Kabupaten/Kota Di Indonesia Berdasarkan Faktor Penyebab Stunting Pada Balita Menggunakan Algoritma K-Means. *JIKO (Jurnal Informatika dan Komputer)*, 6(2), 223-230.
- Faujia, R. A., Setianingsih, E. S., & Pratiwi, H. (2022). Analisis Kluster K-Means Dan Agglomerative Nesting Pada Indikator Stunting Balita Di Indonesia. In *Seminar Nasional Official Statistics (Vol. 2022, No. 1, pp. 1249-1258)*.
- Matdoan, M. Y., Risdiana, F. Y., & Haumahu, G. (2022). Application of the K-Means Cluster for the Classification of Disadvantaged Districts/Cities in Maluku Province. *JRST (Jurnal Riset Sains dan Teknologi)*, 6(1), 61-64.
- Nainggolan, R., & Purba, E. N. (2019). Perbaikan Performa Cluster K-Means Menggunakan Sum Squared Error (SSE) Pada Analisis Online Customer Review Terhadap Produk Toko Online. *Jurnal Times*, 8(2), 1-8.
- Sari, P. F., Pardede, A. M., & Maulita, Y. (2021, June). Pengelompokan Populasi Hewan Ternak Menggunakan Metode Clustering (Studi Kasus: Dinas Pertanian dan Ketahanan Pangan Kabupaten Langkat). In *Seminar Nasional Informatika (SENATIKA)* (pp. 37-46).
- Wibowo, A. P., Darmawan, W., & Amalia, N. (2022). Komparasi Metode Naïve Bayes dan K-Nearest Neighbor terhadap Analisis Sentimen Pengguna Aplikasi Pedulilindungi. *IC-Tech*, 17(1), 18-23.
- Yulia, N., Saragih, R., & Ambarita, I. (2021). Data Mining Pengelompokan Anak Stunting Berdasarkan Usia, Penyebab dan Pekerjaan Orang Tua Dengan Menggunakan Metode Clustering (Studi Kasus: Dinas Kesehatan Kabupaten Langkat). In *Seminar Nasional Informatika (SENATIKA)* (pp. 295-306).