



Penerapan Algoritma K-Means Untuk Clusterisasi Kasus Stunting Di Provinsi Riau

Iqbal Iskandar^a, Deny Jollyta^{b*}

^aInstitut Bisnis dan Teknologi Pelita Indonesia, Jl. Jend. Ahmad Yani No. 78-88 Pekanbaru, iqbal.iskandar@student.pelitaindonesia.ac.id

^bInstitut Bisnis dan Teknologi Pelita Indonesia, Jl. Jend. Ahmad Yani No. 78-88 Pekanbaru, deny.jollyta@lecturer.pelitaindonesia.ac.id

INFORMASI ARTIKEL

Sejarah Artikel:

Diterima Redaksi: 27 Maret 2024

Revisi Akhir: 29 April 2024

Diterbitkan Online: 30 April 2024

KATA KUNCI

Stunting, Clustering, Optimalisasi, Elbow, SSE

KORESPONDENSI

E-mail: deny.jollyta@lecturer.pelitaindonesia.ac.id

A B S T R A C T

Stunting termasuk rencana prioritas untuk ditangani oleh pemerintah Provinsi Riau. Permasalahan yang muncul adalah ketersediaan dan pengolahan data yang belum maksimal pada Dinas Komunikasi Informatika dan Statistik Provinsi Riau. Penelitian ini bertujuan untuk melakukan analisis data *stunting* dengan teknik pengelompokan dan optimalisasi hasil pengelompokan untuk memudahkan dalam mengetahui data *stunting* dan merekomendasikan penanganannya berdasarkan kriteria yang ada. Metode yang digunakan dalam penelitian ini adalah *Clustering* melalui algoritma *K-Means* dan teknik evaluasi *cluster Elbow* dengan memanfaatkan nilai *Sum of Square Error* (SSE). Data *stunting* yang diolah sebanyak 1500 data dari 12 Kabupaten di Provinsi Riau pada Bulan Juli 2023. Data dikelompokkan menggunakan 5 c uji. Hasil pengolahan data memperlihatkan bahwa *cluster* optimal terdapat pada c uji = 3 dengan nilai SSE sebesar 37543,156232997, dimana anggota kelompok pertama terdiri dari 641, kelompok kedua memiliki 594 anggota dan kelompok ketiga terdiri dari 265 anggota.

1. PENDAHULUAN

Dalam penanganan terkait salah satu faktor untuk meningkatkan kesejahteraan masyarakat yaitu, terkait tata kelola data *stunting*. Dinas Komunikasi Informatika dan Statistik Provinsi Riau telah melakukan upaya pengolahan data *stunting*. Tetapi dalam proses pengolahan data tersebut masih belum maksimal dan memerlukan waktu yang relatif lama. Sehingga penelitian ini bertujuan untuk menghasilkan informasi yang akan di sebarluaskan kepada masyarakat dan dapat dijadikan sebagai solusi penyampaian informasi yang akurat terutama tentang kasus *stunting*. *Stunting* itu sendiri merupakan suatu keadaan terganggunya pertumbuhan anak, seperti terganggunya perkembangan otak maupun gangguan pertumbuhan fisik [1]. Dalam jangka panjang *stunting* akan mengakibatkan penurunan kemampuan *kognitif*, penurunan prestasi belajar, penurunan kekebalan tubuh dan beresiko mengalami *obesitas* [2].

Faktor-faktor lain penyebab *stunting* ada dua yaitu, secara langsung seperti asupan gizi dan penyakit infeksi, hipertensi, tidak ASI eksklusif. Kemudian faktor tidak langsungnya adalah pelayanan kesehatan, sosial budaya dan sanitasi lingkungan, status ekonomi, status gizi ibu hamil dan jarak kelahiran yang pendek [3]. Dalam penelitian tentang identifikasi risiko *stunting*

pada anak-anak dengan metode *k-means clustering* menggunakan dataset *kaggle*, berupa cakupan data Provinsi di Indonesia tahun

2016-2018 yaitu, identifikasi kelompok risiko rentan terhadap *stunting* lebih mudah menggunakan metode *clustering*, karena dapat mengetahui faktor-faktor penyebab *stunting* dan cara efektif pencegahan *stunting* pada anak-anak [4].

Penelitian lainnya dengan judul *Regional Mapping in Bangkalan District Based On Potensial Indicators of Total Stunting Using K-Mode Cluster Algorithm* tentang pemetaan wilayah menurut faktor-faktor yang menyebabkan *stunting* menggunakan konsep analisis besarnya koefisien korelasi antar variabel $>0,05$, korelasi *parsial* harus kecil, korelasi antar variabel, ekstraksi dan analisis *cluster* untuk memperoleh hasil akhir tentang data *stunting* [5].

Dari permasalahan tata kelola data *stunting* dan meninjau dari 2 (dua) penelitian sebelumnya, peneliti menerapkan proses *data mining* menggunakan metode *clusterisasi* algoritma *k-means* dalam mengelompokkan data-data balita (pada Bulan Juli tahun 2023) yang ada di 12 Kabupaten Provinsi Riau, dengan 13 kriteria dan optimalisasi hasil *cluster* dengan *elbow method* serta melakukan perbandingan nilai *Sum of Square Error* (SSE) paling besar menggunakan bahasa pemrograman *PHP*.

2. TINJAUAN PUSTAKA

2.1. Data Mining

Data mining dapat di artikan sebagai proses produksi data dalam jumlah besar atau kolonialisme dari hasil suara masyarakat dan adanya pemanfaatan teknologi untuk memudahkan proses seperti halnya kecerdasan buatan [6], yang sinkronisasi dengan konsep *Knowledge Discovery in Database (KDD)*, KDD merupakan kumpulan proses atau langkah-langkah dalam menganalisis data dan berupa hasil keluaran informasi [7].

2.1.1. Clustering

Clustering merupakan algoritma yang termasuk dalam bagian *unsupervised learning* yang tidak menggunakan penamaan variabel pada masing-masing kelompok [8]. *Clustering* juga dapat di artikan proses pemilahan kumpulan data menjadi beberapa kelompok [9], yang mempunyai variasi, kebutuhan data yang sesuai, mengelompokkan data dengan partisi kriteria data yang mempunyai hubungan antara kriteria satu dengan kriteria lainnya [10]. Berfungsi untuk mencari kelompok atribut-atribut ke dalam suatu segmentasi [11].

2.1.2. Algoritma K-Means

Algoritma *k-means* adalah algoritma mengutamakan konsep pengelompokan dari suatu data sesuai kriteria dan ketentuan kelompok, mempunyai perbedaan antara satu data dan data yang lainnya [12] dan adanya relasi antara satu kolom dengan kolom yang lainnya [13]. Algoritma *k-means* adalah algoritma *cluster non hierarchical*, digunakan pada data yang memiliki *outlier*, dengan rata-rata (*mean*) menjadi pusat pengelompokannya [14]. Tujuan pengelompokkan menggunakan algoritma *k-means* yaitu, meminimalkan fungsi variasi objek yang di set dan memaksimalkan variasi data antar kelompok [15]. Penggunaan rumus dalam algoritma ini menggunakan, *euclidean distance* untuk menghitung jarak terdekat antara setiap *centroid*, *euclidean distance* termasuk rumus jarak dalam menentukan kemiripan data dengan keterangan nilai kuantitatif, berikut rumus *euclidean distance*: [16]

$$d(x,y) = |x-y| = \sqrt{\sum (x_i - y_i)^2} \quad (2.1)$$

Dimana d = jarak antara x, y
 x = data pusat *cluster*
 y = data atribut
 i = masing-masing data
 n = jumlah data
 x_i = data pusat *cluster* ke- i
 y_i = data pada masing-masing data ke- i .

2.2. Stunting

Stunting adalah gangguan pertumbuhan anak karena kekurangan asupan gizi, terjangkit infeksi dan stimulasi yang kurang memadai. mengurangi kemungkinan terdampak *stunting* dengan memenuhi kebutuhan konsumsi gizi yang seimbang pada saat mengandung dan pada saat tahap bayi menyusui dengan rutin konsultasi kepada ahli/dokter [17]. Kebutuhan gizi pada balita juga dipengaruhi oleh pengetahuan ibu tentang status gizi balita/anak yang baik akan memberikan dampak pada pertumbuhan balita/anak tersebut [18].

2.3. Elbow Method dan Sum of Square Error

Elbow method merupakan metode untuk optimalisasi *cluster* dengan cara melihat grafik, berupa garis yang paling membentuk siku dan penurunan yang paling besar antara masing-masing titik *cluster*. Melalui perbandingan/selisih terbesar dari nilai *Sum Of Square Error (SSE)* setiap c uji dan dijadikan sebagai hasil evaluasi c uji yang optimal dengan rumus SSE seperti dibawah ini: [19]

$$SSE = \sum_k^k = I \sum x_i e s_k // |x_i - c_k| / 2 \quad (2.2)$$

Dimana k = banyaknya *cluster*

c = *cluster* ke- k

x_i = data yang berada di masing-masing *cluster*

2.4. Website

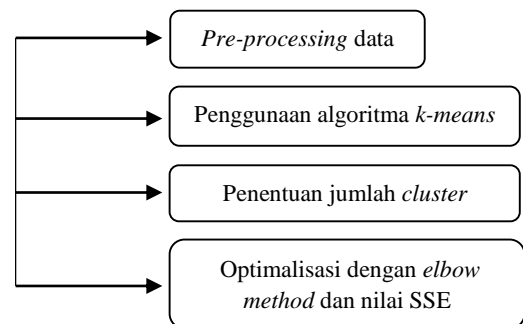
Website adalah sebuah wadah teknologi atau sarana yang berfungsi untuk menampung data dan informasi ditampilkan, dengan pola dan konsep sesuai kebutuhan *user* [20]. Berdasarkan kegunaannya *website* bisa di manfaatkan sebagai *platform* untuk keperluan pribadi, promosi bahkan wadah penyalur informasi. Dan umumnya *website* terbagi atas 2 (dua) kategori yaitu; *website statis* dan *website dinamis* [21].

2.5. PHP

Dulunya *PHP* dikenal dengan nama *Form Interpreted* yang merupakan kumpulan *skrip* yang berfungsi untuk mengolah data dari formulir *website* [22]. *PHP* mempunyai kelebihan sebagai bahasa pemrograman *multi platform* diberbagai mesin dan sistem informasi, yang dapat dipakai secara gratis (*open source*) dan dapat dijalankan secara *runtime* melalui *console* [23].

3. METODOLOGI

Dalam menyusun dan menentukan *clusterisasi (c)* yang relevan, di bawah ini ada beberapa tahapan seperti pada gambar 1 [24]



Gambar 1. Tahapan Pengolahan Data

Pengolahan data diawali dengan ketentuan *pre-processing* data untuk membersihkan data-data yang kurang lengkap dan akan diproses menggunakan *clusterisasi* algoritma *k-means* dengan jumlah data yang siap di olah berjumlah 1500 data. Ada beberapa tahapan dalam pengelompokkan menggunakan algoritma *k-means*, yaitu: [25]

1. Tentukan jumlah nilai *cluster* yang akan di bentuk.
2. Tentukan posisi *centroid* setiap *cluster (c)* secara *random*
3. Hitung jarak terpendek dengan *euclidean distance*.
4. Kelompokkan data jarak terpendek pada tahap ke-3

5. Tentukan titik *centroid* baru dengan mencari rata-rata dari data *centroid* yang sama.
6. Lakukan perbandingan hasil jarak terdekat *centroid* baru dengan hasil jarak terdekat *centroid* sebelumnya.
7. Jika hasil pada tahapan sama maka perhitungan selesai dan jika hasilnya berbeda perhitungan dilanjutkan.

Kemudian dilakukan optimalisasi dengan *elbow method* dan melihat selisih nilai SSE yang paling besar sehingga didapat informasi dari data *stunting* yang sudah diolah.

4. HASIL DAN PEMBAHASAN

Dari dataset yang telah dikelola dan telah dilakukan *filterisasi*, dilakukan pengujian manual menggunakan *microsoft excel* dengan jumlah data sebanyak 100 baris data dan data berjumlah 1500 data dilakukan pengolahan menggunakan sistem berbasis *website* dengan setiap kriteria/variabelnya dapat dilihat pada tabel 1 di bawah ini.

Tabel 1 : Kriteria/Variabel Data *Stunting*

No.	Kriteria/Variabel	Keterangan
1.	Jenis Kelamin	<i>Gender</i> balita yang melakukan pengecekan prosedur <i>stunting</i>
2.	Berat Badan Lahir	Berat Badan bayi yang di timbang setelah lahir
3.	Tinggi Badan Lahir	Tinggi Badan bayi yang di ukur setelah lahir
4.	Berat Badan Pengecekan	Berat Badan balita pada saat pengecekan <i>stunting</i>
5.	Tinggi Badan Pengecekan	Tinggi Badan balita pada saat pengecekan <i>stunting</i>
6.	LiLA	Hasil pengukuran Lingkar Lengan Atas balita
7.	BB/U	Berat Badan berdasarkan usia
8.	TB/U	Tinggi Badan berdasarkan usia
9.	BB/TB	Berat Badan menurut Tinggi Badan
10.	ZS BB/U	Penilaian status gizi berdasarkan standar WHO pada Berat Badan berdasarkan usia
11.	ZS TB/U	Penilaian status gizi berdasarkan standar WHO pada Tinggi Badan berdasarkan usia
12.	ZS BB/TB	Penilaian status gizi berdasarkan standar WHO pada Berat Badan menurut Tinggi Badan
13.	Kenaikan Berat Badan	Perubahan Berat Badan saat pengecekan <i>stunting</i>

(Sumber: Dinas Komunikasi Informatika dan Statistik Provinsi Riau, 2023)

Kemudian tabel 2 di bawah ini merupakan bentuk normalisasi atau konversi variabel yang digunakan dalam proses perhitungan nantinya, karena sebagai salah satu syarat algoritma *k-means* yaitu dataset yang menjadi keputusan harus berupa data *numerik*.

Tabel 2 : Deskripsi Konversi Variabel

Variabel	Deskripsi	Konversi
Jenis	Laki-Laki	1
Kelamin	Perempuan	2
BB/U	Normal	1
	Kurang	2
	Sangat kurang	3
	Risiko Lebih	4
TB/U	Pendek	1
	Sangat pendek	2
BB/TB	Gizi Baik	1
	Gizi Kurang	2
	Gizi Buruk	3
	Risiko Gizi Lebih	4
	Gizi Lebih	5
	Obesitas	6
Naik	N Naik	1
Berat Badan	T Tidak Naik	2
	O Tidak ada data timbang bulan lalu	3

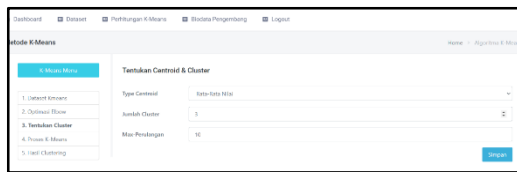
Kemudian analisa ini melibatkan identifikasi data *statistik deskriptif* yang dapat mengkategorikan melalui nilai minimum, maksimum, rata-rata dan *Standar Deviation* (SD). Dan uji *multikorelinearitas* merupakan uji yang berfungsi untuk pengecekan apakah ditemukan korelasi antar masing-masing variabel dari dataset uji. Jika ada korelasi maka ditandai dengan H_1 dan jika tidak ada maka H_0 , dengan tingkat nilai value signifikansi $= >5\%$ (H_0) dan $<5\%$ (H_1) dengan SPSS.

Selanjutnya tahap pengolahan menggunakan sistem yang menunjukkan data berhasil di *upload* dalam sistem, dimana untuk tampilannya dapat dilihat seperti pada gambar 2 di bawah ini:

nama balita	usia	berat	tinggi	bb_u	tb_u	bb_tb	zs_bb_u	zs_tb_u	zs_bb_tb	kenaikan			
nama balita 1	2	5.5	50	5.4	45.5	15	3	-0.04	2	-0.09	1	-0.2	3
nama balita 2	2	1.9	40	2.4	47.6	11	1	-0.05	1	-0.04	1	-0.28	3
nama balita 3	2	3.2	47	3.3	51.5	15	3	-0.44	1	-0.8	2	-0.24	2
nama balita 4	1	3.2	50	3.6	50.7	15.2	2	-0.68	2	-0.2	1	-0.51	2
nama balita 5	2	3.2	49	3.5	47	15	1	-0.61	2	-0.01	1	-0.66	3
nama balita 6	1	3	47	3.6	50.8	17	2	-0.21	1	-0.05	1	-0.01	3

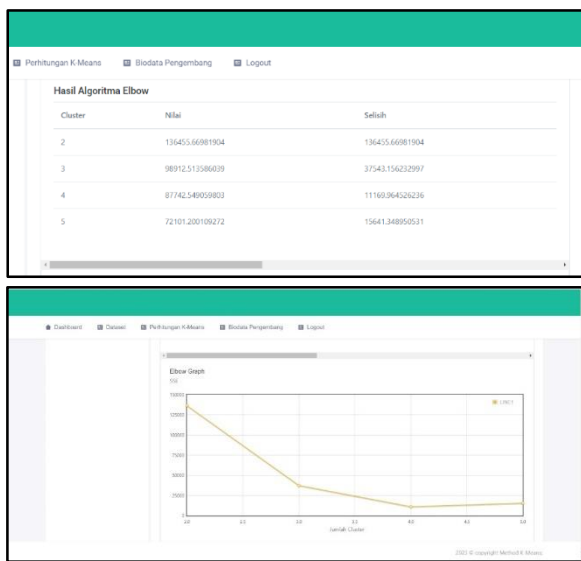
Gambar 2. Dataset

Tahap berikutnya penentuan jumlah *cluster* yang digunakan sebanyak 3 *cluster*, dengan proses perulangan 10 kali, dengan proses pemilihan dilakukan secara *random*, yang ditampilkan pada gambar 3 penentuan jumlah *cluster* di bawah ini. Kemudian untuk penjelasan terkait penentuan jumlah *cluster* sebanyak 3 *cluster* dapat dilihat pada tahap berikutnya yaitu optimalisasi.



Gambar 3. Penentuan Jumlah Cluster

Pada tahap optimalisasi ini menggunakan *elbow method* yang menunjukkan hasil yang paling akurat dengan grafik yang menampilkan garis membentuk siku dan mengalami penurunan yang paling besar terletak pada 3 cluster berdasarkan c uji = 5 dengan nilai selisih SSE yang paling besar terletak pada cluster 3 yaitu $136455,66981904-98912,513586039= 37543,156232997$, sehingga penentuan jumlah cluster = 3 dapat diterima, untuk lebih jelasnya dapat dilihat pada gambar 4 di bawah ini.



Gambar 4. Optimalisasi Cluster

Selanjutnya pada gambar 5 riwayat iterasi, hasil yang diperoleh perulangan ke-10 yang memenuhi, karena pada perulangan ke-10 hasil jarak terdekat antara perulangan ke-10 dengan perulangan sebelumnya (ke-9) di dapatkan nilai yang sama. Sehingga berdasarkan ketentuan algoritma *k-means*, perhitungan dapat di stop/diberhentikan apabila pada saat proses hitung data memperoleh hasil yang sama.

Gambar 5. Riwayat Iterasi/Perulangan

Tahap terakhir berupa hasil proses hitung dari riwayat iterasi didapat pada tabel di bawah ini, yaitu masing-masing pembagian data balita yang terindikasi *stunting* berdasarkan 3 kelompok (*cluster* 1-3) yang berjumlah 1500 data.

Tabel 3 : Hasil Cluster Optimal C Uji =3

Cluster	Jumlah
1	641
2	594
3	265
Total	1500

Dari hasil proses perhitungan dataset yang berjumlah 1500 balita yang terindikasi *stunting* pada Bulan Juli 2023 di Provinsi Riau menunjukkan pada masing-masing kelompok pada tabel 3 hasil cluster, yaitu:

Interpretasi/ pengetahuan yang diperoleh adalah:

1. Pengaruh yang signifikan dalam penentuan balita/anak yang terdampak *stunting* di antara 13 variabel dari dataset uji yang telah dilakukan adalah rata-rata dari; Tinggi Badan saat melakukan pengukuran, dipengaruhi oleh Tinggi Badan Lahir Berat Badan penilaian status gizi berdasarkan standar WHO pada Tinggi Badan dan Berat Badan berdasarkan usia, penilaian status gizi berdasarkan standar WHO pada Tinggi Badan dan Berat Badan.
2. Nilai signifikansi BB Lahir, TB Lahir, Berat Badan, Tinggi Badan, Lingkar Lengan atas, penilaian status gizi berdasarkan standar WHO pada Berat Badan berdasarkan usia, Tinggi badan berdasarkan usia, penilaian status gizi berdasarkan standar WHO pada Tinggi Badan berdasarkan usia dan penilaian status gizi berdasarkan standar WHO pada Berat Badan dan Tinggi Badan berdasarkan usia adalah ($,000 < 0,05$) sehingga penilaian yang menentukan penyebab *stunting* dapat diterima dibuktikan juga dengan uji *multikolinearitas* dimana variabel-variabel tersebut mempunyai korelasi dalam penentu *stunting*.
3. Rata-rata berat badan lahir balita pada Provinsi Riau relatif normal, namun yang menjadi faktor lain adalah kenaikan berat badan sehingga, tinggi badan dengan rata-rata 1,23 yang berarti melebihi tinggi badan dengan keterangan pendek.
4. Rata-rata jumlah terindikasi *stunting* pada bulan Juli dengan dataset 1500 pada berat badan menurut usia paling tertinggi adalah cluster 1 (1,6677), rata-rata menengah adalah cluster 2 (1,6323), cluster 3 (1,6615), termasuk BB/U dengan risiko berat badan lebih (*underweight*). Tetapi walaupun cluster dengan jumlah tertinggi, tingkat kenaikan berat badan pada anak cukup tinggi juga.
5. Anggota cluster terdiri dari beberapa balita yaitu; C1: 641. C2: 594. C3: 265 balita, kemudian yang menjadi salah satu faktor utama bahwa balita/anak yang terindikasi *stunting* adalah penilaian status gizi berdasarkan standar WHO pada Tinggi Badan berdasarkan usia dengan nilai negatif atau di bawah rata-rata dan nilai rata-rata positif tertinggi adalah pada variabel berat badan.
6. Cluster 1 balita secara keseluruhan dengan tinggi badan menurut usia ada keterangan pendek berjumlah:456 balita, sangat pendek berjumlah:185 balita. Cluster 1 untuk balita dengan jenis kelamin laki-laki berjumlah 374 balita, naik

- 144, tidak naik 145, tidak ada data timbang bulan lalu 85. Balita laki-laki dengan tinggi badan menurut usia keterangan sangat pendek berjumlah:102 balita dengan berat badan normal ada 27 balita, berat badan kurang ada 43 balita, berat badan sangat kurang 32 balita. Gizi baik 74 balita, gizi buruk 1 balita, gizi kurang 16 balita, risiko gizi lebih 6 balita, gizi lebih 4 balita dan *obesitas* 1 balita. Selanjutnya balita dengan jenis kelamin laki-laki keterangan pendek ada 272 balita, dengan berat badan normal ada 161 balita berat badan kurang ada 88 balita, berat badan sangat kurang 22 balita dan risiko berat badan lebih 1 balita. Gizi baik 211 balita, gizi buruk 7 balita, gizi kurang 30 balita, risiko gizi lebih 15 balita, gizi lebih 6 balita, dan *obesitas* 3 balita. *Cluster 1* untuk balita dengan jenis kelamin perempuan berjumlah 267 balita, kenaikan data timbang 103 balita, tidak naik 111 balita, tidak ada data timbang bulan lalu 53 balita. Balita laki-laki dengan tinggi badan menurut usia keterangan sangat pendek berjumlah:83 balita dengan berat badan normal ada 24 balita, berat badan kurang ada 25 balita, berat badan sangat kurang 34 balita. Gizi baik 57 balita, gizi buruk 6 balita, gizi kurang 12 balita, risiko gizi lebih 4 balita, gizi lebih 2 balita dan *obesitas* 2 balita. Selanjutnya balita dengan jenis kelamin perempuan keterangan pendek ada 184 balita, dengan berat badan normal ada 104 balita berat badan kurang ada 66 balita, berat badan sangat kurang 12 balita dan risiko berat badan lebih 2 balita. Gizi baik 144 balita, gizi buruk 3 balita, gizi kurang 21 balita, risiko gizi lebih 12 balita, gizi lebih 3 balita, dan *obesitas* 1 balita.
7. *Cluster 2* secara keseluruhan dengan tinggi badan menurut usia ada keterangan pendek berjumlah:519 balita, sangat pendek berjumlah:75 balita. *Cluster 2* untuk balita dengan jenis kelamin laki-laki berjumlah 327 balita, kenaikan data timbang 145 balita, tidak naik 114 balita, tidak ada data timbang bulan lalu 68 balita. Balita laki-laki dengan tinggi badan menurut usia keterangan sangat pendek berjumlah:38 balita dengan berat badan normal ada 8 balita, berat badan kurang ada 15 balita, berat badan sangat kurang 14 balita dan risiko berat badan lebih 1 balita. Gizi baik 33 balita, gizi buruk 1 balita, gizi kurang 2 balita, risiko gizi lebih 2 balita. Selanjutnya balita dengan jenis kelamin laki-laki keterangan pendek ada 289 balita, dengan berat badan normal ada 161 balita berat badan kurang ada 119 balita, berat badan sangat kurang 9 balita. Gizi baik 218 balita, gizi buruk 7 balita, gizi kurang 38 balita, risiko gizi lebih 19 balita, gizi lebih 6 balita, dan *obesitas* 1 balita. *Cluster 2* untuk balita dengan jenis kelamin perempuan berjumlah 267 balita, kenaikan data timbang 105 balita, tidak naik 117 balita, tidak ada data timbang bulan lalu 45 balita. Balita laki-laki dengan tinggi badan menurut usia keterangan sangat pendek berjumlah:37 balita dengan berat badan normal ada 10 balita, berat badan kurang ada 12 balita, berat badan sangat kurang 15 balita. Gizi baik 23 balita, gizi buruk 1 balita, gizi kurang 9 balita, risiko gizi lebih 2 balita, gizi lebih 1 balita dan *obesitas* 1 balita. Selanjutnya balita dengan jenis kelamin perempuan keterangan pendek ada 230 balita, dengan berat badan normal ada 102 balita, berat badan kurang ada 108 balita, berat badan sangat kurang 18 balita dan risiko berat badan lebih 2 balita. Gizi baik 179 balita, gizi buruk 4 balita, gizi kurang 27 balita, risiko gizi lebih 12 balita, gizi lebih 5 balita, dan *obesitas* 3 balita.
8. *Cluster 3* balita secara keseluruhan dengan tinggi badan menurut usia ada keterangan pendek berjumlah:183 balita, sangat pendek berjumlah:82 balita. *Cluster 3* untuk balita dengan jenis kelamin laki-laki berjumlah 147 balita, kenaikan data timbang 50 balita, tidak naik 59 balita, tidak ada data timbang bulan lalu 38 balita. Balita laki-laki dengan tinggi badan menurut usia keterangan sangat pendek berjumlah:48 balita dengan berat badan normal ada 11 balita, berat badan kurang ada 19 balita, berat badan sangat kurang 18 balita. gizi baik 36 balita, gizi kurang 7 balita, risiko gizi lebih 4 balita dan gizi lebih 1 balita. Selanjutnya balita dengan jenis kelamin laki-laki keterangan pendek ada 99 balita, dengan berat badan normal ada 57 balita berat badan kurang ada 33 balita, berat badan sangat kurang 9 balita. Gizi baik 74 balita, gizi kurang 11 balita, risiko gizi lebih 4 balita. *Cluster 3* untuk balita dengan jenis kelamin perempuan berjumlah 118 balita, kenaikan data timbang 23 balita, tidak naik 36 balita, tidak ada data timbang bulan lalu 25 balita. Balita laki-laki dengan tinggi badan menurut usia keterangan sangat pendek berjumlah:34 balita dengan berat badan normal ada 14 balita, berat badan kurang ada 12 balita, berat badan sangat kurang 15 balita, risiko berat badan lebih 1 balita. Gizi baik 23 balita, gizi buruk 2 balita, gizi kurang 4 balita, risiko gizi lebih 4 balita dan *obesitas* 1 balita. Selanjutnya balita dengan jenis kelamin perempuan keterangan pendek ada 84 balita, dengan berat badan normal ada 57 balita, berat badan kurang ada 19 balita, berat badan sangat kurang 5 balita dan risiko berat badan lebih 3 balita. Gizi baik 59 balita, gizi buruk 2 balita, gizi kurang 10 balita, risiko gizi lebih 7 balita, gizi lebih 4 balita, dan *obesitas* 2 balita.

5. KESIMPULAN DAN SARAN

Sebagai bentuk implementasi dari metode (algoritma *k-means*) yang digunakan dalam proses perhitungan data *stunting* khususnya di Provinsi Riau hasil data *stunting* pada bulan Juli 2023 adalah rata-rata relatif sangat kurang dengan jumlah balita jenis kelamin laki-laki lebih banyak jika di dibandingkan dengan jumlah balita jenis kelamin perempuan yaitu; untuk kelompok 1 dengan jumlah 641 balita terhadap 1 kriteria, untuk kelompok 2 dengan jumlah 594 balita terhadap 7 kriteria, dan kelompok 3 dengan jumlah 265 balita terhadap 5 kriteria. Kemudian kriteria yang menjadi faktor utama balita/anak terindikasi *stunting* adalah tinggi badan yang relatif kurang dari masing-masing *cluster* dan pengaruh dari status gizi balita yang relatif tidak terkontrol sehingga balit/anak tersebut kekurangan gizi bahkan gizi berlebih. Sebagai bentuk pengembangan dapat dilakukan dengan menggunakan optimalisasi metode yang lainnya seperti metode *Davies Bouldin Index* (DBI) dan *Silhouette Index* (SI).

UCAPAN TERIMA KASIH.

Terimakasih kepada Bapak/Ibu Dosen, terutama Dosen Pembimbing yang telah memberikan arahan, bantuan selama proses penelitian ini, terimakasih juga kepada kedua orang tua, keluarga dan orang-orang terdekat saya, yang selalu memberikan *support* terbaik sehingga dapat menyelesaikan penelitian ini.

DAFTAR PUSTAKA

- [1] S. A. Sakti, "Pengaruh Stunting pada Tumbuh Kembang Anak Periode Golden Age," *J. Ilm. Fak. Kegur. dan Ilmu Pendidik.*, vol. 6, no. 1, pp. 169–175, 2020, [Online]. Available: <http://ejournal.unsub.ac.id/index.php/FKIP>
- [2] D. A. Muharyani, Putri Widita, "Aplikasi Strategi Intervensi Simulation Game Dalam Upaya Pencegahan Stunting Pada Anak," *J. Pengabd. Sriwij.*, vol. 7, no. 2, pp. 789–794, 2019, doi: 10.37061/jps.v7i2.9771.
- [3] A. Ramdhani, H. Handayani, and A. Setiawan, "Hubungan Pengetahuan Ibu Dengan Kejadian Stunting," *Semnas Lppm*, vol. ISBN: 978-, pp. 28–35, 2020, [Online]. Available: <https://semnaslppm.ump.ac.id/index.php/semnaslppm/article/view/122/117>
- [4] D. Setiawan, Rudiansyah, and S. R. Fadila, "Identifikasi Faktor Risiko Stunting Pada Anak-Anak Dengan Metode K Means Clustering," *J. Tek. Inform.*, vol. 3, no. 1, pp. 1–7, 2023, [Online]. Available: <http://teknologipintar.org/index.php/teknologipintar/article/view/103/78>
- [5] A. N. Sari, D. H. Qurotu'ain, F. Harianto, and S. Z. Jannah, "Regional Mapping in Bangkalan District Based on Potential Indicators of Total Stunting Using K-Mode Cluster Algorithm," *Media Gizi Indones. (Natl. Nutr. Journal)*, no. 1, pp. 76–82, 2022, doi: 10.20473/mgi.v17i1sp.76-82.
- [6] J. S. Roberts and L. N. Montoya, "In Consideration of Indigenous Data Sovereignty: Data Mining as a Colonial Practice," *Futur. Technol. Conf.*, p. 12, 2023, [Online]. Available: <https://arxiv.org/ftp/arxiv/papers/2309/2309.10215.pdf>
- [7] Y. Irawan, "Penerapan Data Mining Untuk Evaluasi Data Penjualan Menggunakan Metode Clustering Dan Algoritma Hirarki Divisive," *JTIULM*, vol. 04, no. 1, pp. 1–8, 2019, [Online]. Available: <http://www.jtiulm.ti.ft.uilm.ac.id/index.php/jtiulm/article/view/34/33>
- [8] H. Priyatman, F. Sajid, and D. Haldivany, "Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa," *JEPIN*, vol. 5, no. 1, pp. 62–66, 2019, [Online]. Available: <https://core.ac.uk/download/pdf/294888966.pdf>
- [9] M. Cahyanti, M. R. D. Septian, E. R. Swedia, and R. A. Salim, "Pengelompokan Citra Kendaraan (Motor Dan Mobil) Berdasarkan Bentuk Menggunakan Algoritma K-Means," *Sebatik*, vol. 22, no. 2, pp. 153–160, 2018, doi: 10.46984/sebatik.v22i2.322.
- [10] D. Jollyta, Prihandoko, A. Hajjah, E. Haerani, and M. Siddik, *Algoritma Klasifikasi Untuk Pemula Solusi Python Dan Rapidminer*, Pertama. Yogyakarta: Deepublish Publisher (grup Penerbitan CV BUDI UTAMA), 2023.
- [11] A. Rifa'i, G. G. Setiaji, and V. Vydia, "Penggunaan Metode K-Means Pada Analisa Dan Klasifikasi Capres 2019 Di Twitter," *J. Pengemb. Rekayasa dan Teknol.*, vol. 15, no. 1, p. 43, 2019, doi: 10.26623/jprt.v15i1.1489.
- [12] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan," *J. Tekno Kompak*, vol. 15, no. 2, pp. 25–36, 2021, doi: 10.33365/jtk.v15i2.1162.
- [13] T. Hardiani, "Analisis Clustering Kasus Covid 19 di Indonesia Menggunakan Algoritma K-Means," *J. Nas. Pendidik. Tek. Inform.*, vol. 11, no. 2, pp. 156–165, 2022, doi: 10.23887/janapati.v11i2.45376.
- [14] P. M. Arumsari, U. Hayati, and G. Dwilestari, "Implementasi Pada Pengelompokan Data Stunting Balita Menggunakan Algoritma Clustering K-Medoids," *J. Ilm. Betrik Baemah Teknol. Infromasi dan Komput.*, vol. 14, no. 01, pp. 166–174, 2023, [Online]. Available: https://scholar.google.com/scholar?cluster=5943007357213103313&hl=id&as_sdt=0,5&as_ylo=2023
- [15] D. Kusbianto, I. A. Mashudi, and F. D. Prianggoro, "Implementasi Metode Rapid Centroid Estimation Dan K-Means untuk Klasterisasi Sekolah Berdasarkan Fasilitas (Studi Kasus Sekolah Negeri Kabupaten Malang)," *Semin. Inform. Apl. Polinema*, pp. 254–259, 2020, [Online]. Available: <http://jurnalti.polinema.ac.id/index.php/SIAP/article/view/780/231>
- [16] M. Nishom, "Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square," *J. Inform. J. Pengemb. IT*, vol. 4, no. 1, pp. 20–24, 2019, doi: 10.30591/jpit.v4i1.1253.
- [17] Kemkes, "Kementrian kesehatan Direktorat Promosi Kesehatan dan Pemberdayaan Masyarakat, Pencegahan stunting pada anak," 2019. <https://promkes.kemkes.go.id/pencegahan-stunting>
- [18] Y. B. Prasetyo, P. Permatasari, and H. D. Susanti, "The effect of mothers' nutritional education and knowledge on children's nutritional status: a systematic review," *Int. J. Child Care Educ. Policy*, vol. 17, no. 1, pp. 1–16, 2023, doi: 10.1186/s40723-023-00114-7.
- [19] D. Jollyta, M. Siddik, H. Mawengkang, and S. Efendi, *Teknik Evaluasi Cluster Solusi Menggunakan Python dan RapidMiner*, Pertama. Yogyakarta: Deepublish (Grup Penerbitan CV BUDI UTAMA), 2021.
- [20] E. R. Rahmi, E. Yumami, and N. Hidayasari, "Analisis Metode Pengembangan Sistem Informasi Berbasis Website: Systematic Literature Review," *Remik, Ris. dan E-Jurnal Manaj. Inform. Komput.*, vol. 7, no. 1, pp. 821–834, 2023, doi: 10.33395/remik.v7i1.12177.
- [21] T. Limbong and Sriadhi, *Pemrograman Web Dasar*, Cetakan 1. Yayasan Kita Menulis, 2021. [Online]. Available: <http://digilib.unimed.ac.id/48203/1/Book.pdf>
- [22] M. Wali *et al.*, "Pengantar 15 Bahasa Pemrograman Terbaik Di Masa Depan (Refrensi & Coding Untuk Pemula)," 1st ed., Effitra, Ed. Jambi: PT. SONpedia Publishing Indonesia, 2023. [Online]. Available: https://books.google.co.id/books?hl=id&lr=&id=uIWxEAAAQBAJ&oi=fnd&pg=PP1&dq=bahasa+pemrograman+php&ots=c6zpjKfuLj&sig=6IQwTI1k17q17JJPJ-hjIZw4PnCU&redir_esc=y#v=onepage&q=bahasa+pemrograman+php&f=false
- [23] B. Satria and A. Rahim, "Pengembangan Aplikasi Monitoring Keterlambatan Kehadiran Siswa SMK Muhammadiyah 1 Penajam Paser Utara Berbasis Web Development of Web-Based Student Attendance Monitoring Application of SMK Muhammadiyah 1 Penajam Paser Utara," vol. 4, no. 1, 2024, [Online]. Available: <https://jurnal-stiepari.ac.id/index.php/safari/article/view/1152/1068>
- [24] D. Jollyta, S. Efendi, M. Zarlis, and H. Mawengkang, "Optimasi Cluster Pada Data Stunting: Teknik Evaluasi Cluster Sum of Square Error dan Davies Bouldin Index," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, pp. 918–926, 2019, doi: 10.30645/senaris.v1i0.100.
- [25] H. Jamaludin and B. Y. Dhamahita, "K-Means Clustering Analysis on the Distribution of Stunting Cases in Mojokerto Regency in June 2022," *J. Media Pratama J. Media Pratama*, vol. 15, no. 1, pp. 34–43, 2023, [Online]. Available: <https://jurnal.inf.co.id/index.php/jurnalmediapratama/article/view/320/237>