

ANALISIS CLUSTERING PROVINSI BERDASARKAN PREVALENSI STUNTING BALITA MENGGUNAKAN ALGORITMA SINGLE DAN COMPLETE LINKAGE

Septian Wulandari

Program Studi Teknik Informatika, Universitas Indraprasta PGRI
Jl Nangka Raya No 58C, Tanjung Barat, Kec. Jagakarsa, Jakarta Selatan
septian.pmb09@rocketmail.com

ABSTRAK

Gizi adalah salah satu parameter dalam menilai keberhasilan pembangunan kesehatan di suatu negara untuk menciptakan sumber daya manusia yang bermutu. Stunting merupakan gangguan dimana laju pertumbuhan pada anak sehingga pertumbuhan tinggi badan dan berat badan pada anak terhambat akibat kekurangan gizi. Balita yang menderita stunting tidak hanya mengalami gangguan pertumbuhan pada fisik, melainkan dapat mengakibatkan anak mudah terserang penyakit, gangguan perkembangan pada otak dan kecerdasan yang mengakibatkan menurunnya kualitas sumber daya manusia di Indonesia. Frekuensi prevalensi stunting di Indonesia naik setiap tahunnya terbukti pada tahun 2019 prevalensi stunting pada balita mencapai 35,6% dan pada tahun 2020 sebesar 36,8%. Oleh karena itu, permasalahan stunting merupakan permasalahan yang serius dan pemerintah Indonesia harus segera menangani masalah ini. Penanganan stunting yang cepat dan tepat dapat memperlambat pertumbuhan stunting. Data yang diperoleh melalui BPS pada tahun 2018-2022 ditemukan balita yang mengalami gizi buruk ataupun tinggi badan tidak sesuai dengan usianya dan belum adanya penanganan seperti *clustering* provinsi yang tepat untuk menanggulangi stunting dan dapat menekan laju pertumbuhan stunting yang kian meningkat. Algoritma yang digunakan pada penelitian ini adalah *single linkage* dan *complete linkage* dengan data 34 provinsi di Indonesia dan 11 indikator yang menjadi acuan. Hasil pada penelitian ini menghasilkan 2 *cluster* dengan *cluster* 1 beranggotakan 6 provinsi dan *cluster* 2 beranggotakan 24 provinsi. Algoritma *complete linkage* menjadi algoritma terbaik dibandingkan algoritma *single linkage* dengan nilai korelasi *chopenetic* sebesar 0,8608356.

Kata Kunci: *Stunting, Single Linkage, Complete Linkage*

ABSTRACT

Nutrition is one of the parameters in assessing the success of health development in a country to create quality human resources. Stunting is a disorder in which the growth rate of children results in stunted growth in height and weight due to malnutrition. Toddlers who suffer from stunting not only experience physical growth disorders, but can result in children being susceptible to disease, impaired development of the brain and intelligence which results in a decline in the quality of human resources in Indonesia. The frequency of stunting prevalence in Indonesia increases every year as evidenced by the fact that in 2019 the prevalence of stunting in children under five reached 35.6% and in 2020 it was 36.8%. Therefore, the problem of stunting is a serious problem and the Indonesian government must immediately address this problem. Fast and appropriate handling of stunting can slow the growth of stunting. Data obtained through BPS in 2018-2022 found that toddlers were suffering from poor nutrition or their height was not appropriate for their age and there was no appropriate treatment such as provincial clustering to tackle stunting and could reduce the increasing growth rate of stunting. The algorithm used in this research is single linkage and complete linkage with data from 34 provinces in Indonesia and 11 indicators as references. The results of this study produced 2 clusters with cluster 1 consisting of 6 provinces and cluster 2 consisting of 24 provinces. The complete linkage algorithm is the best algorithm compared to the single linkage algorithm with a *chopenetic* correlation value of 0.8608356.

Key Word: *Stunting, Single Linkage, Complete Linkage*

PENDAHULUAN

Stunting merupakan gangguan laju pertumbuhan pada anak sehingga pertumbuhan tinggi badan dan berat badan pada anak terhambat tidak sesuai pada usianya. Ciri-ciri balita yang mengalami stunting terlihat pada tinggi badan yang lebih rendah dibandingkan rata-rata pertumbuhan

pada anak atau dapat direpresentasikan yaitu nilai *z-score* pada tinggi badan menurut umur (TB/U) minus 2 dari standar deviasi (SD) (Ni'mah et al., 2015). Hal ini disebabkan oleh kurangnya asupan gizi pada anak dalam waktu yang cukup lama, asupan ASI eksklusif, kehamilan *preterm*, memberikan makanan

yang tidak optimal, kekurangan gizi, pendidikan, sanitasi pada lingkungan. Frekuensi prevalensi stunting di Indonesia naik setiap tahunnya terbukti pada tahun 2019 prevalensi stunting pada balita mencapai 35,6% dan pada tahun 2020 sebesar 36,8% (Apriyani et al., 2023). Pada tahun 2020, menurut Asian Development Bank (ADB) kasus prevalensi anak di Indonesia yang menderita stunting pada usia di bawah lima tahun (balita) menduduki peringkat kedua tercatat sebesar 31,8% dan kasus tertinggi terletak pada negara Timor Leste yaitu sebesar 48,8% (Indah Syafitri Nasution & Susilawati, 2022).

Tumbuh kembang pada balita adalah fase pertumbuhan yang sangat rawan terhadap lingkungan maka pada fase inilah diperlukan perhatian khusus seperti kecukupan gizi pada balita. Gizi adalah salah satu parameter dalam menilai keberhasilan pembangunan kesehatan di suatu negara untuk menciptakan sumber daya manusia yang bermutu (Rosari & Rini, 2013). Permasalahan gizi seperti *stunting* yang dialami balita dapat memperlambat perkembangan anak, sehingga efek negatif yang berlangsung pada kehidupan selanjutnya dengan gejala penurunan pada intelektual, mudah terserang penyakit yang tidak menular, penurunan pada produktivitas bahkan dapat mengakibatkan kemiskinan serta risiko melahirkan bayi dengan berat badan lahir yang rendah (Ni'mah et al., 2015). Selain itu, balita yang menderita stunting tidak hanya mengalami gangguan pertumbuhan pada fisik saja, melainkan dapat mengakibatkan anak mudah terserang penyakit, serta gangguan perkembangan pada otak dan kecerdasan yang mengakibatkan menurunnya kualitas sumber daya manusia di Indonesia (Rahman et al., 2023).

Oleh karena itu, permasalahan stunting merupakan permasalahan yang serius dan pemerintah Indonesia harus segera menangani masalah ini. Penanganan stunting yang cepat dan tepat dapat memperlambat pertumbuhan stunting. Langkah awal yang harus dilakukan pemerintah adalah dengan melakukan regulasi yaitu menjadikan kasus prevalensi stunting menjadi prioritas utama dalam memberantas stunting pada balita di kalangan masyarakat. Penilaian kasus stunting pada balita ditentukan pada pengukuran anatomi tubuh manusia yang dikenal dengan nama

anthropometri (tinggi badan ideal berdasarkan usia) seperti Umur (U), Berat Badan (BB), dan Tinggi Badan (TB) (Apriyani et al., 2023). Agar kebijakan dalam menangani kasus prevalensi stunting pada balita yang dilakukan oleh pemerintah Indonesia tepat sasaran maka perlu dilakukan *clustering* pada balita yang mengalami stunting dengan menentukan dari beberapa aspek yang saling berhubungan seperti data usia pada balita, status gizi, dan penyebab stunting (Ranjawali et al., 2023). *Clustering* dilakukan berdasarkan beberapa indikator seperti indeks penanganan stunting, bayi usia kurang dari 6 bulan yang mendapatkan asi eksklusif, presentase balita pendek dan sangat pendek, presentase provinsi yang mencapai 80% imunisasi dasar lengkap pada bayi, presentase balita obesitas (bb/tb) kelompok umur 0-59 bulan, prevalensi balita gizi buruk (0-23 bulan) dan (0-59 bulan), prevalensi balita gizi kurang (0-23 bulan) dan (0-59 bulan), serta prevalensi balita kekurangan gizi (0-23 bulan) dan (0-59 bulan).

Analisis *cluster* adalah salah satu alat statistik untuk menganalisis data pada kasus prevalensi stunting balita. Analisis *cluster* adalah metode yang diaplikasikan untuk *clustering* objek ke dalam *cluster* sesuai dengan informasi yang dijumpai pada data (Tri et al., 2019). Hasil pada *clustering* dikatakan sesuai jika terdapat data yang homogenitas dalam satu *cluster* dan terdapat heterogenitas antar satu *cluster* dengan *cluster* lainnya. Terdapat dua jenis metode dalam *clustering* yaitu *hierarki clustering* dan *non hierarki clustering*. Metode *agglomerative clustering* merupakan salah satu metode dari *hierarki clustering*. *Agglomerative clustering* dimulai dengan menginput *cluster* tunggal (berisikan satu objek per *cluster*) pada tingkat bawah kemudian menggabungkan dua *cluster* serta membangun hirarki secara *bottom-up* dari *cluster* (Wulandari, 2023).

Pada metode *agglomerative clustering* memiliki empat algoritma yaitu algoritma *complete linkage*, *ward's*, *single linkage*, dan *average linkage*. Pada algoritma *single linkage* memiliki keunggulan yaitu hampir setiap *cluster* yang dihasilkan pada setiap kelompok selalu mirip dengan anggota lainnya (Handoyo et al., 2014). Sedangkan algoritma *complete linkage* menggunakan jarak maksimal yaitu jarak antara satu *cluster*

dengan *cluster* lain diukur berdasarkan objek data yang memiliki jarak yang paling jauh (Ramadhani et al., 2018).

Terdapat penelitian mengenai clusterisasi stunting pada balita yang dilakukan oleh Putri Apriyani pada tahun 2023 dengan menerapkan algoritma k-means di Desa Tegalwangi dengan hasil terdapat dua *cluster* dengan parameter yang digunakan yaitu umur, berat badan, dan tinggi badan (Apriyani et al., 2023). Penelitian lainnya juga dilakukan oleh Ratna Dewi pada tahun 2023 dengan menerapkan metode *fuzzy subtractive clustering* pada wilayah kerawanan stunting dan diperoleh hasil terdapat tujuh *cluster* kecamatan (Christyanti et al., 2023). Selanjutnya terdapat penelitian yang dilakukan oleh Andrea Tri Rian pada tahun 2019 dengan menerapkan metode *hierarchal clustering* menggunakan metode *agglomerative* pada runtun waktu dan diperoleh dua *cluster* kabupaten/kota di Provinsi Kalimantan Timur (Tri et al., 2019). Untuk menangani kasus stunting pada balita di Indonesia perlu dilakukan *clustering* dengan data prevalensi stunting balita di Indonesia sehingga diharapkan dapat menunjukkan pengetahuan dan informasi untuk pemerintahan Indonesia untuk menerapkan usulan-usulan yang tepat dalam menangani prevalensi stunting pada balita. Sehingga, pada penelitian dilakukan penelitian yaitu analisis *clustering* provinsi berdasarkan prevalensi stunting balita menggunakan algoritma *single* dan *complete linkage*.

METODE PENELITIAN

Pendekatan yang dilakukan pada penelitian ini yaitu menggunakan studi literatur deskriptif kuantitatif yaitu mengumpulkan referensi kepustakaan yang membantu dalam keberhasilan penelitian ini. Sedangkan, pada fase deskriptif kuantitatif dilakukan dengan menginput data, menyelidiki data, dan menafsirkan data sesuai yang diperlukan pada penelitian ini. Fase pertama yang dilakukan yaitu dengan mengumpulkan data yang didapatkan pada data BPS dan diunduh melalui laman <https://www.bps.go.id>. Data diambil pada tahun 2018 hingga 2022 yaitu data prevalensi stunting balita pada 34 Provinsi yang ada di Indonesia. Terdapat 11 indikator data yang digunakan yaitu:

- x_1 : Indeks khusus pada penanganan stunting
- x_2 : Presentasi bayi usia < 6 bulan yang memperoleh asi eksklusif
- x_3 : Presentasi balita berbadan pendek dan sangat pendek
- x_4 : Presentasi provinsi dengan 80% imunisasi dasar lengkap sejak bayi
- x_5 : Presentasi balita dengan berat badan obesitas (bb/tb) pada kelompok umur 0-59 bulan
- x_6 : Prevalensi balita dengan gizi buruk pada kelompok umur (0-59 bulan)
- x_7 : Prevalensi balita dengan gizi buruk pada kelompok umur (0-23 bulan)
- x_8 : Prevalensi balita dengan gizi kurang pada kelompok umur (0-59 bulan)
- x_9 : Prevalensi balita dengan gizi kurang pada kelompok umur (0-23 bulan)
- x_{10} : Prevalensi balita dengan kekurangan gizi pada kelompok umur (0-59 bulan)
- x_{11} : Prevalensi balita dengan kekurangan gizi pada kelompok umur (0-23 bulan)

Dalam menganalisis *cluster* perlu dilakukan uji multikolinieritas yaitu hubungan linier antara semua variable. Multikolinieritas dapat diketahui dengan menghitung nilai *Variance Inflation Factor (VIF)*. Jika nilai *VIF* >10 maka dapat dikatakan terjadi multikolinieritas antara variabel prediktor. Berikut merupakan rumus untuk menghitung nilai *VIF* (Nurissaidah Ulinnuh, 2020):

$$VIF = \frac{1}{1-R^2} \quad (1)$$

Adanya multikolinieritas dapat dihitung dengan mencari nilai *Tolerance* dengan rumus:

$$Tolerance = \frac{1}{VIF} \quad (2)$$

Jika nilai *Tolerance* > 0,10 maka tidak ada multikolinieritas. Jika terdapat adanya multikolinieritas maka dapat diatasi dengan mengeluarkan variable yang tidak memiliki hubungan atau yg tidak memiliki keterkaitan. *Kaiser Meyer Olkin (KMO)* digunakan untuk membandingkan besaran koefisien korelasi dan besaran koefisien parsial. Kemudian, Uji *Bartlett Sphericity* digunakan untuk menguji interdependensi antar variabel yang merupakan indikator pada suatu faktor. Analisis ini digunakan untuk mengetahui variabel yang tidak memiliki korelasi antara satu dengan lainnya pada populasi (Edo Verdian, 2019). Untuk menghitung *KMO*

dapat dihitung dengan menggunakan rumus (Sauddin, n.d.):

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p a_{ij}^2} \quad (3)$$

dengan:

i : 1,2,3, ..., p

j : 1,2,3, ..., p , dengan $j \neq i$

r_{ij}^2 : koefisien korelasi sederhana antara variable i dan j

a_{ij}^2 : koefisien korelasi parsial antara variable i dan j

Jika nilai uji $KMO > 0,5$ maka dapat dikatakan bahwa data yang dianalisis sudah menangkup kecukupan data yang diolah. Sedangkan, untuk statistic *Chi-square* dapat dihitung dengan rumus (Nurul Afida & Edy Sulistiyawan, 2014):

$$\chi_{hitung}^2 = \sum_{i=1}^k \frac{O_i - E_i^2}{E_i} \quad (4)$$

dengan:

χ^2 : Nilai *Chi Square*

O_i : Frekuensi pada hasil pengamatan

k : jumlah. Nilai interval

Pengambilan keputusan didasarkan pada tolak H_0 jika $\chi_{hitung}^2 > \chi^2$.

Metode yang digunakan pada penelitian ini adalah metode *agglomerative clustering* dengan algoritma *single* dan *complete linkage*. *Single linkage* adalah salah satu algoritma dari metode *agglomerative clustering* yang berdasarkan pada jarak paling kecil antar objek. Algoritma ini dimulai dengan memilih jarak yang kecil pada matriks $D = \{d_{ij}\}$ dan menggabungkan objek yang memiliki kemiripan seperti objek U dan V untuk mendapatkan *cluster* (UV). Selanjutnya yaitu dengan menghitung W yaitu jarak antara *cluster* (UV) dan *cluster* lainnya yang dihitung dengan menggunakan rumus (Tri et al., 2019):

$$d_{(UV)W} = \min(d_{UV}, d_{VW}) \quad (5)$$

dengan:

d_{UV} : jarak antar tetangga terdekat pada cluster U dan W

d_{VW} : jarak antar tetangga terdekat pada cluster V dan W

Sedangkan, algoritma *complete linkage* atau dikenal juga dengan algoritma pautan lengkap merupakan kebalikan dari algoritma *single linkage*. Algoritma *complete linkage* merupakan mengukur antar *cluster* yang satu dengan *cluster* lainnya mengacu pada objek

yang mempunyai jarak yang terjauh. Objek-objek pada setiap cluster disambungkan antara satu dan lainnya dengan kesamaan jarak maksimum maupun minimum. Tahapan awal algoritma *complete linkage* adalah menghitung nilai minimum dari $D = d_{ij}$ dan digabungkan dengan objek lainnya yang memiliki kesamaan seperti U dan V untuk mendapatkan *cluster* (UV), sedangkan jarak antar cluster lain W dan (UV) dihitung menggunakan rumus (Nurissaidah Ulinnuh, 2020):

$$d_{(UV)W} = \max(d_{UV}, d_{VW}) \quad (6)$$

Untuk mengevaluasi *cluster* menggunakan *Silhouette Coefficient* yaitu untuk mengetahui kualitas serta kekuatan pada *cluster*, dan mengevaluasi suatu objek yang ditempatkan pada suatu *cluster* (Handoyo et al., 2014). Berikut ini adalah tahapan dalam menghitung *Silhouette Coefficient*:

1. Menghitung jarak rata-rata dari suatu objek i dengan objek lain yang terdapat pada satu cluster:

$$a(i) = \frac{1}{|A|-1} \sum_{j \in A, j \neq i} d(i, j) \quad (7)$$

Dengan j adalah objek lain pada satu cluster A dan $d(i, j)$ adalah jarak antara objek i dan j .

2. Menghitung jarak rata-rata dari objek i dengan semua objek di cluster lainnya, kemudian ambil nilai yang paling kecil dengan rumus:

$$d(i, C) = \frac{1}{|A|} \sum_{j \in C} d(i, j) \quad (8)$$

Dengan $d(i, C)$ merupakan jarak rata-rata objek i dengan semua objek pada cluster C dengan $A \neq C$.

3. Menghitung nilai *Silhouette Coefficient* dengan rumus:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (9)$$

Berikut adalah daftar penilaian *Silhouette Coefficient* dapat dilihat pada Tabel 1.

Tabel 1. Ukuran nilai silhouette coefficient

| Skala | Keterangan |
|------------------------|--------------------------------|
| $0,7 < s(i) \leq 1$ | <i>Strong Structure Medium</i> |
| $0,5 < s(i) \leq 0,7$ | <i>Medium Structure</i> |
| $0,25 < s(i) \leq 0,5$ | <i>Weak Structure</i> |
| $s(i) \leq 0,25$ | <i>No Structure</i> |

HASIL DAN PEMBAHASAN

Analisis statistik deksriptif dilakukan untuk mengetahui indikator prevalensi stunting pada balita apakah dapat digunakan dengan baik atau tidak. Indikator prevalensi stunting pada

balita terdiri dari indeks penanganan stunting, bayi usia kurang dari 6 bulan yang mendapatkan asi eksklusif, presentase balita pendek dan sangat pendek, presentase provinsi yang mencapai 80% imunisasi dasar lengkap pada bayi, presentase balita obesitas (bb/tb) kelompok umur 0-59 bulan, prevalensi balita gizi buruk (0-23 bulan dan (0-59 bulan), prevalensi balita gizi kurang (0-23 bulan dan (0-59 bulan), serta prevalensi balita kekurangan gizi (0-23 bulan dan (0-59 bulan). Berikut adalah tabel analisis statistik deskriptif yang sudah dilakukan normalisasi dari setiap indikator yang diukur dengan Mean, Nilai Min, Median, dan Nilai Max dapat dilihat pada Tabel 2.

Tabel 2. Statistik deskriptif variabel penelitian

| Variabel | Mean | Min | Median | Max |
|----------|--------|--------|--------|--------|
| x_1 | 0,6384 | 0,5079 | 0,6407 | 0,7989 |
| x_2 | 0,4529 | 0,2550 | 0,4421 | 0,6419 |
| x_3 | 0,0709 | 0 | 0,0862 | 0,1599 |
| x_4 | 0,7072 | 0,0431 | 0,7799 | 1 |
| x_5 | 0,0741 | 0,0306 | 0,0762 | 0,1298 |
| x_6 | 0,042 | 0,0175 | 0,0421 | 0,0717 |
| x_7 | 0,0415 | 0,0115 | 0,0401 | 0,1008 |
| x_8 | 0,1455 | 0,0957 | 0,1384 | 0,2201 |
| x_9 | 0,1207 | 0,0717 | 0,1208 | 0,1739 |
| x_{10} | 0,1901 | 0,278 | 0,1754 | 0,2932 |
| x_{11} | 0,1647 | 0,1038 | 0,1679 | 0,2431 |

Berdasarkan tabel 2 menunjukkan bahwa Presentasi provinsi dengan 80% imunisasi dasar lengkap sejak bayi tergolong tinggi sebesar 100% berada pada provinsi Jambi, Kepulauan Riau, DKI Jakarta, Jawa Tengah, DI Yogyakarta, Bali, dan Nusa Tenggara Barat. Untuk Presentasi balita berbadan pendek dan sangat pendek cukup rendah sebesar 0,31% berada pada provinsi Kalimantan Utara, dan indeks khusus pada penanganan stunting tinggi pada provinsi Nusa Tenggara Barat yaitu sebesar 72,97%. Dalam menganalisis *cluster* perlu dilakukan uji multikolinieritas. Multikolinieritas dapat diketahui dengan melakukan perhitungan nilai *Varians Inflation Factor (VIF)* dan hasil perhitungan dapat dilihat pada Tabel 3.

Tabel 3. Hasil perhitungan nilai vif dan tolenrance

| Variabel bebas | Collinearity Statistics | |
|----------------|-------------------------|-------|
| | Tolerance | VIF |
| x_1 | 0,723 | 1,383 |
| x_2 | 0,778 | 1,285 |
| x_3 | 0,654 | 1,528 |
| x_4 | 0,675 | 1,482 |

| | | |
|----------|-------|-------|
| x_5 | 0,943 | 1,061 |
| x_6 | 0,550 | 1,818 |
| x_7 | 0,751 | 1,332 |
| x_8 | 0,642 | 1,559 |
| x_9 | 0,731 | 1,367 |
| x_{10} | 0,585 | 1,711 |
| x_{11} | 0,884 | 1,132 |

Jika nilai *VIF* >10 maka dapat dikatakan terjadi multikolinieritas antara variable prediktor dan jika nilai *Tolerance* > 0,10 maka tidak ada multikolinieritas.. Terlihat pada Tabel 3 bahwa semua nilai *VIF* < 10 serta nilai *Tolerance* > 0,10 artinya tidak terdapat multikolinieritas antara variable prediktor. Selanjutnya dilakukan uji *Kaiser Meyer Olkin (KMO)* yang digunakan untuk membandingkan besaran koefisien korelasi dan besaran koefisien parsial. Kemudian, Uji Bartlett Sphercity digunakan untuk menguji interdependensi antar variabel yang merupakan indikator pada suatu faktor. Hasil uji *KMO* dan uji *Bartlett Sphercity* dapat dilihat pada Tabel 4.

Tabel 4. Hasil uji bartlett sphercity

| | KMO | | |
|---------------------------|-------------------------------|------------|---------|
| | Measure of Sampling Adequacy. | Chi-Square | p-value |
| <i>Bartlett Sphercity</i> | 0,523 | 69,789 | 0,351 |

Berdasarkan table 4 dengan uji *Kaiser Meyer Olkin (KMO)* diperoleh nilai uji *KMO* 0,523 > 0,5 artinya yang dianalisis sudah menangkup kecukupan data yang diolah. Sedangkan uji Bartlett Sphercity diperoleh nilai *p-value* sebesar 0,351 > 0,05 artinya varians pada data homogen sehingga asumsi untuk homogenitas sudah terpenuhi. Hubungan korelasi antar variabel juga dapat dilihat melalui mariks korelasi pada Gambar 1.

Terlihat pada Gambar 1 bahwa semakin mendekati angka1 atau -1 maka semakin kuat korelasi antar setiap variable, angka positif menunjukkan korelasi positif dan angka negatif menunjukkan korelasi negatif. Sedangkan, angka 0 tidak terdapat korelasi antar kedua variable.



Gambar 1. Matriks korelasi antar variabel

Prose selanjutnya yaitu melakukan *clustering* provinsi prevalensi stunting pada balita menggunakan algoritma *single* dan *complete linkage*. Proses *clustering* dilakukan menggunakan aplikasi R Studio dan diperoleh nilai korelasi *chopenetic* pada Tabel 5.

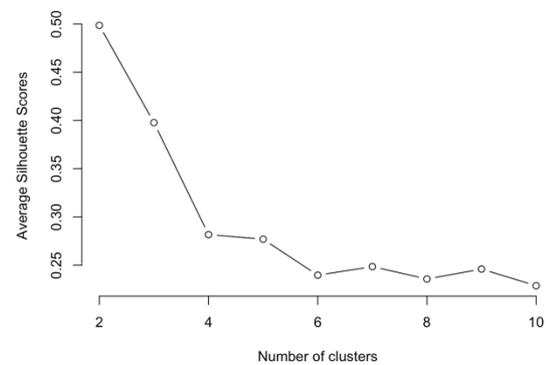
| Algoritma | Korelasi <i>Chopenetic</i> |
|------------------|----------------------------|
| Single Linkage | 0,5463915 |
| Complete Linkage | 0,8608356 |

Berdasarkan Table 5 menunjukkan bahwa nilai pada korelasi *chopenetic* angka tertinggi terletak pada algoritma *complete linkage* dengan nilai 0,8608356 sehingga algoritma *complete linkage* dipilih pada penelitian ini. Kemudian, Langkah selanjutnya adalah melakukan clustering dimulai dengan $k = 2$ sampai dengan 9 dengan menghitung nilai *Silhouette Coefficient* dan didapatkan hasil pada Tabel 6 dan grafiknya dapat dilihat pada Gambar 2.

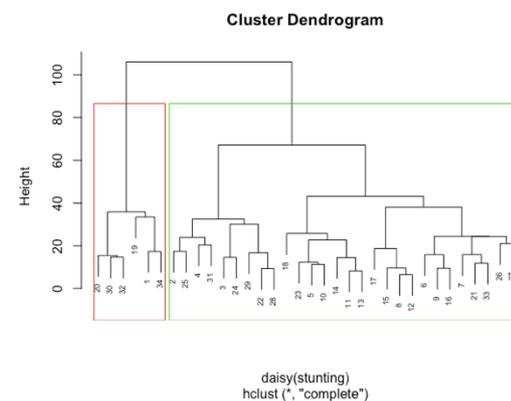
| Jumlah cluster | Nilai <i>Silhouette Coefficient</i> |
|----------------|-------------------------------------|
| 2 | 0,4986312 |
| 3 | 0,3977930 |
| 4 | 0,2816879 |
| 5 | 0,2396488 |
| 6 | 0,2485166 |
| 7 | 0,2357174 |
| 8 | 0,2459782 |
| 9 | 0,2287614 |

Tabel 6 dan Gambar 2 menunjukkan bahwa angka tertinggi perhitungan nilai *Silhouette Coefficient* terletak pada jumlah cluster 2 dengan nilai 0,4986312 yang masuk ke dalam kategori *weak structure*. Hal ini berarti hasil cluster yang diperoleh memiliki ikatan yang lemah antar objek. Sehingga pada penelitian

ini dilakukan *clustering* dengan jumlah cluster yaitu 2 cluster dan dihasilkan cluster pada Gambar 3.



Gambar 2. Grafik perhitungan nilai *silhouette coefficient*



Gambar 3. Dendrogram hasil *clustering* algoritma *complete linkage*

Untuk memudahkan dalam menganalisis hasil clustering maka perlu dilakukan labeling pada setiap provinsi yang ada di Indonesia. Pelabelan provinsi yang ada di Indonesia dapat dilihat pada Tabel 7.

| No | Provinsi | No | Provinsi |
|----|----------------------|----|---------------------|
| 1 | Aceh | 18 | Nusa Tenggara Barat |
| 2 | Sumatera Utara | 19 | Nusa Tenggara Timur |
| 3 | Sumatera Barat | 20 | Kalimantan Barat |
| 4 | Riau | 21 | Kalimantan Tengah |
| 5 | Jambi | 22 | Kalimantan Selatan |
| 6 | Sumatera Selatan | 23 | Kalimantan Timur |
| 7 | Bengkulu | 24 | Kalimantan Utara |
| 8 | Lampung | 25 | Sulawesi Utara |
| 9 | Kep. Bangka Belitung | 26 | Sulawesi Tengah |

| | | | |
|----|---------------|----|-------------------|
| 10 | Kep. Riau | 27 | Sulawesi Selatan |
| 11 | Dki Jakarta | 28 | Sulawesi Tenggara |
| 12 | Jawa Barat | 29 | Gorontalo |
| 13 | Jawa Tengah | 30 | Sulawesi Barat |
| 14 | DI Yogyakarta | 31 | Maluku |
| 15 | Jawa Timur | 32 | Maluku Utara |
| 16 | Banten | 33 | Papua Barat |
| 17 | Bali | 34 | Papua |

Gambar 3 menunjukkan dendrogram hasil *clustering* pada algoritma *complete linkage*. Terlihat bahwa terdapat 2 *cluster* yang dihasilkan dengan *cluster* 1 berjumlah 6 provinsi dan *cluster* 2 berjumlah 28 provinsi dan daftar anggota pada masing-masing cluster terlihat pada Tabel 8.

Tabel 8. Hasil anggota cluster

| Cluster | Provinsi | Total |
|---------|---|-------|
| 1 | Kalimantan Barat, Sulawesi Barat, Maluku Utara, Nusa Tenggara Timur, Aceh, dan Papua | 6 |
| 2 | Sumatera Utara, Sulawesi Utara, Riau, Maluku, Sumatera Barat, Kalimantan Utara, Gorontalo, Bali, Sulawesi Tenggara, Nusa Tenggara Barat, Kalimantan Timur, Jambi, Kep. Riau, DI Yogyakarta, DKI Jakarta, Jawa Tengah, Bali, Jawa Timur, Lampung, Jawa Barat, Sumatera Selatan, Kep. Bangka Belitung, Banten, Bengkulu, Kalimantan Tengah, Papua Barat, Sulawesi Tengah, dan Sulawesi Tenggara | 24 |

Penyebaran prevalensi stunting pada balita di Indonesia akan berkurang jika indeks penanganan stunting bertambah, presentasi bayi usia < 6 bulan yang memperoleh asi eksklusif, Presentasi balita berbadan pendek dan sangat pendek berkurang, presentasi provinsi dengan 80% imunisasi dasar lengkap sejak bayi semakin bertambah, presentasi balita dengan berat badan obesitas (bb/tb) pada kelompok umur 0-59 bulan berkurang, prevalensi balita dengan gizi buruk pada kelompok umur (0-23 bulan) dan (0-59 bulan) berkurang, prevalensi balita dengan gizi kurang pada kelompok umur (0-23 bulan) dan (0-59 bulan) berkurang, dan prevalensi balita dengan kekurangan gizi pada kelompok umur (0-23 bulan) dan (0-59 bulan) berkurang. Berdasarkan uraian diatas, dapat disimpulkan bahwa cluster 1 merupakan provinsi dengan tingkat prevalensi balita dengan gizi buruk pada kelompok umur (0-23 bulan) dan (0-59 bulan) yang tinggi, prevalensi balita dengan

gizi kurang pada kelompok umur (0-23 bulan) dan (0-59 bulan) yang tinggi, dan prevalensi balita dengan kekurangan gizi pada kelompok umur (0-23 bulan) dan (0-59 bulan) yang tinggi. Anggota provinsi pada *cluster* 1 adalah Kalimantan Barat, Sulawesi Barat, Maluku Utara, Nusa Tenggara Timur, Aceh, dan Papua. Sedangkan, *cluster* 2 merupakan provinsi dengan indeks khusus pada penanganan stunting yang tinggi, Presentasi bayi usia < 6 bulan yang memperoleh asi eksklusif yang tinggi, Presentasi balita berbadan pendek dan sangat pendek yang rendah, Presentasi provinsi dengan 80% imunisasi dasar lengkap sejak bayi yang tinggi, serta Presentasi balita dengan berat badan obesitas (bb/tb) pada kelompok umur 0-59 bulan yang rendah. Anggota provinsi pada cluster 2 adalah Sumatera Utara, Sulawesi Utara, Riau, Maluku, Sumatera Barat, Kalimantan Utara, Gorontalo, Bali, Sulawesi Tenggara, Nusa Tenggara Barat, Kalimantan Timur, Jambi, Kep. Riau, DI Yogyakarta, DKI Jakarta, Jawa Tengah, Bali, Jawa Timur, Lampung, Jawa Barat, Sumatera Selatan, Kep. Bangka Belitung, Banten, Bengkulu, Kalimantan Tengah, Papua Barat, Sulawesi Tengah, dan Sulawesi Tenggara.

SIMPULAN DAN SARAN

Kesimpulan yang dapat diambil dari penelitian *clustering* prevalensi stunting pada balita adalah: pertama algoritma *complete linkage* memiliki performa yang lebih baik dibandingkan dengan algoritma *single linkage* dengan nilai korelasi *chopenetic* sebesar 0,8608356, kedua jumlah cluster yang terbaik yang dihasilkan adalah sebanyak 2 cluster dengan menggunakan penilaian *Silhouette Coefficient* dengan nilai 0,4986312 yang masuk ke dalam kategori *weak structure*. Cluster 1 merupakan provinsi yahng mempunyai prevalensi stunting yang tinggi dan Cluster 2 merupakan provinsi yahng mempunyai prevalensi stunting yang rendah.

Saran yang diberikan pada penelitian ini adalah perlu dikembangkan lebih lanjut mengenai pemilihan jumlah cluster tidak hanya menggunakan penilaian *Silhouette Coefficient* tetapi juga bisa menggunakan Metode *Elbow* ataupun *Gap Statistic*. Selain itu, jumlah data yang digunakan dapat dengan data tidak hanya pada tahun 2018 sampai

dengan 2022, namun dapat menambahkan drai tahun-tahun sebelumnya.

DAFTAR PUSTAKA

- Apriyani, P., Dikananda, A. R., & Ali, I. (2023). Penerapan Algoritma K-Means Dalam Klasterisasi Kasus Stunting Balita Desa Tegalwangi. *Hello World Jurnal Ilmu Komputer*, 2(1), 20–33.
- Christyanti, R. D., Sulaiman, D., Utomo, A. P., & Ayyub, M. (2023). Clustering Wilayah Kerawanan Stunting Menggunakan Metode Fuzzy Subtractive Clustering. *Jurnal Ilmiah Teknologi Informasi Asia*, 17(1).
- Edo Verdian. (2019). Analisis Faktor Yang Merupakan Intensi Perpindahan Merek Transportasi Online Di Surabaya. *Agora*, 7(1).
- Handoyo, R., Rumani M, R., & Nasution, S. M. (2014). Perbandingan Metode Clustering Menggunakan Metode Single Linkage Dan K-Means Pada Pengelompokan Dokumen. *Jsm Stmik Mikroskil*, 15(2), 73–82.
- Indah Syafitri Nasution, & Susilawati. (2022). Analisis Faktor Penyebab Kejadian Stunting Pada Balita Usia 0-59 Bulan. *Florona Jurnal Ilmu Kesehatan*, 1(2), 82–87.
- Ni'mah, K., Nadhiroh, S. R., Kesehatan, D. G., & Kesehatan, F. (2015). Faktor Yang Berhubungan Dengan Kejadian Stunting Pada Balita. *Media Gizi Indonesia*, 10(1), 13–19.
- Nurissaidah Ulinuh, R. V. (2020). Analisis Cluster Dalam Pengelompokan Provinsi Di Indonesia Berdasarkan Variabel Penyakit Menular Menggunakan Metode Complete Linkage, Average Linkage Dan Ward. *Infotekjar : Jurnal Nasional Informatika Dan Teknologi Jaringan*, 5(1). <https://doi.org/10.30743/infotekjar.v5i1.2464>
- Nurul Afida, & Edy Sulistiyawan. (2014). Analisis Faktor-Faktor Yang Mempengaruhi Prestasi Belajar Siswa Kelas X Sma Al-Islam Krian. *J Statistika*, 6, 1–6.
- Rahman, H., Rahmah, M., & Nur Saribuan. (2023). Upaya Penanganan Stunting Di Indonesia. *Jurnal Ilmu Pemerintahan Suara Khatulistiwa (Jipsk)*, 8(01), 44–59.
- Ramadhani, L., Purnamasari, I., Deny, F., & Amijaya, T. (2018). Penerapan Metode Complete Linkage Dan Metode Hierarchical Clustering Multiscale Bootstrap (Studi Kasus: Kemiskinan Di Kalimantan Timur Tahun 2016) Application Of Complete Linkage Method And Hierarchical Clustering Multiscale Bootstrap Method (Case Study: Poverty In East Kalimantan Year 2016). *Jurnal Eksponensial*, 9(1).
- Ranjawali, R., Carmen Talakua, A., Thimotius Abineno, R., Wira Wacana Sumba Jruprato No, K., Waingapu, K., Sumba Timur, K., & Tenggara Timu, N. (2023). Clustering Stunting Pada Balita Dengan Metode K-Means Di Puskesmas Kanatang. *Sati: Sustainable Agricultural Technology Innovation*, 80–92.
- Rosari, A., & Rini, A. (2013). Hubungan Diare Dengan Status Gizi Balita Di Kelurahan Lubuk Buaya Kecamatan Koto Tangah Kota Padang. *Jurnal Kesehatan Andalas*, 2(3). <http://jurnal.fk.unand.ac.id>
- Sauddin, A. (N.D.). *Analisis Faktor-Faktor Yang Mempengaruhi Kepatuhan Wajib Pajak Orang Pribadi Dalam Memenuhi Kewajiban Membayar Pajak Kendaraan Bermotor (Pkb)*.
- Tri, A., Dani, R., Wahyuningsih, S., & Rizki, N. A. (2019). Penerapan Hierarchical Clustering Metode Agglomerative Pada Data Runtun Waktu. *Jambura Journal Of Mathematics*, 1. <http://ejournal.ung.ac.id/index.php/jjom,P->
- Wulandari, S. (2023). *Clustering Indonesian Provinces On Prevalence Of Stunting Toddlers Using Agglomerative Hierarchical Clustering*. 16(2), 197–276. <https://doi.org/10.30998/faktorexacta.v16i2.17186>