

I. PENDAHULUAN

1.1. Latar Belakang

Kementerian Kesehatan mengungkapkan bahwa stunting merupakan salah satu masalah gizi utama di Indonesia (Kementerian Kesehatan RI, 2023a). Stunting merupakan gangguan pertumbuhan dan perkembangan pada balita yang diakibatkan oleh asupan gizi yang buruk dan infeksi kronis berulang, yang ditandai dengan tinggi badan yang berada di bawah standar usianya, yaitu tinggi badan balita lebih rendah dari -2 standar deviasi dibawah standar ketetapan pada kurva pertumbuhan anak World Health Organization (WHO). Kondisi ini dapat bersifat permanen apabila asupan nutrisi tidak mencukupi dan infeksi kronis berlangsung selama 1000 Hari Pertama Kehidupan (HPK) balita (WHO, 2015). Dampak dari kondisi stunting meliputi pertumbuhan fisik yang terhambat, gangguan perkembangan kognitif dan motorik, ukuran fisik tubuh yang tidak optimal, serta peningkatan risiko penyakit kronis yang berhubungan dengan gizi di masa dewasa (Ariani, 2020).

Kementerian Kesehatan mengumumkan hasil Survei Kesehatan Indonesia (SKI) yang menunjukkan adanya penurunan prevalensi stunting dari tahun ke tahun. Namun, prevalensi stunting di Indonesia pada tahun 2023 masih tergolong tinggi yaitu mencapai 21,5% (Kementerian Kesehatan RI, 2024a). Berdasarkan standar WHO, prevalensi stunting yang baik di suatu negara harus berada kurang dari 20% (Kementerian Kesehatan RI, 2023b). Oleh karena itu, angka stunting di Indonesia masih membutuhkan perhatian serius oleh pemerintah.

Sejalan dengan komitmen Indonesia untuk terus mengurangi angka stunting, sebagaimana diatur dalam Peraturan Presiden No. 72 Tahun 2021 tentang percepatan penurunan stunting, pemerintah telah merancang Program Bangga Kencana dan 8 aksi konvergensi. Untuk mendukung upaya pemerintah dalam penurunan stunting, penelitian ini akan menerapkan teknik *clustering* untuk mengidentifikasi daerah-daerah rawan stunting di Indonesia berdasarkan beberapa faktor penyebab stunting dengan menggunakan pendekatan *data mining*. *Data mining* adalah proses untuk mengidentifikasi informasi tersembunyi dalam basis data dan merupakan bagian dari *Knowledge Discovery in Databases* (KDD) yang bertujuan menemukan informasi dan pola yang berguna dalam data (Tarigan et al., 2022).

Berdasarkan Kementerian Kesehatan RI, faktor penyebab stunting terbagi menjadi langsung dan tidak langsung. Faktor penyebab stunting secara langsung yaitu masalah kurangnya asupan gizi dan status kesehatan, terutama penyakit

infeksi. Sementara, penyebab tidak langsung mencakup berbagai faktor seperti ketahanan pangan (akses pangan bergizi), lingkungan sosial (pemberian makanan bayi dan anak, kebersihan, pendidikan, dan tempat kerja), lingkungan kesehatan (akses pelayanan preventif dan kuratif), dan lingkungan pemukiman (akses air bersih, air minum, dan sarana sanitasi. (Kementerian Kesehatan RI, 2023a).

Kementerian Dalam Negeri (Kemendagri) melalui Ditjen Bina Pembangunan Daerah (Bangda) melakukan aksi 8 konvergensi untuk penurunan stunting. Salah satu upaya yang dilakukan adalah memberikan pelayanan kesehatan preventif dan kuratif kepada seluruh masyarakat di Indonesia guna mencegah dan menurunkan angka stunting. *Monitoring* data cakupan layanan kesehatan untuk penurunan stunting dan sebaran kasus stunting di Indonesia tersedia di website resmi aksi Bangda Kemendagri. Variabel cakupan layanan ini terdiri dari remaja putri yang mengonsumsi tablet tambah darah (TTD), remaja putri yang menerima layanan pemeriksaan status anemia (Hemoglobin), calon pengantin atau ibu yang menerima tablet tambah darah (TTD), pasangan calon pengantin yang mendapatkan bimbingan perkawinan dengan materi pencegahan stunting, ibu hamil kurang energi kronik (KEK) yang mendapatkan tambahan asupan gizi, ibu hamil yang mengonsumsi tablet tambah darah (TTD), dan lain-lain. Variabel-variabel ini merupakan faktor yang mempengaruhi angka stunting di Indonesia. Oleh karena itu, dalam penelitian ini akan menggunakan variabel cakupan pelayanan esensial dan kasus stunting dari 38 provinsi di Indonesia.

Clustering merupakan salah satu teknik dalam *unsupervised learning* yang digunakan untuk mengelompokkan sejumlah objek ke dalam beberapa *cluster* berdasarkan kemiripan pola yang serupa di *cluster* yang sama dan pola objek yang tidak mirip di *cluster* lainnya (Rotul Muhima et al., 2021). Metode *clustering* terdiri dari *Hierarchical Clustering* dan *Non-Hierarchical (Partitional) Clustering*. *Hierarchical Clustering* adalah metode pengelompokan objek yang memiliki kemiripan yang membentuk struktur hirarki dalam bentuk pohon atau *dendrogram* (Aselnino & Wijayanto, 2024). Algoritma *Hierarchical Clustering* terdiri dari *Agglomerative Hierarchy Clustering* dan *Divisive Hierarchy Clustering*. Sedangkan *Non-Hierarchical Clustering* merupakan metode pengelompokan objek ke beberapa *cluster* secara langsung tanpa membentuk diagram pohon atau *dendrogram* (Aselnino & Wijayanto, 2024). Algoritma *Non-Hierarchical Clustering* terdiri dari K-Means, Fuzzy C-Means, dan K-Medoid.

Berdasarkan hasil penelitian oleh (Aselnino & Wijayanto, 2024), metode K-Means terbukti lebih optimal dibandingkan metode hirarki maupun non-hirarki lainnya, seperti Fuzzy C-Means dan *Hierarchical Clustering* dalam membentuk

cluster provinsi di Indonesia berdasarkan indikator *women empowerment*. Penelitian tersebut menggunakan *Silhouette Index*, *Dunn Index*, dan *Calinski-Harabasz Index* sebagai pengukur validasi. Metode K-Means memiliki nilai tertinggi pada *Silhouette Index* dan *Calinski-Harabasz Index*, yang menunjukkan bahwa metode ini paling optimal dalam membentuk *cluster*.

Tabel 1. Pemilihan Model *Clustering* Terbaik

Metode	<i>Silhouette Index</i>	<i>Dunn Index</i>	<i>Calinski-Harabasz Index</i>
Hierarchical	0,2254	0,2109	10,2234
K-Means	0,2434	0,2116	10,6696
Fuzzy C-Means	0,2272	0,2140	10,1118

Sumber: (Aselnino & Wijayanto, 2024)

Penelitian lain terkait *clustering* menggunakan algoritma K-Means juga telah dilakukan oleh (Handayani & Sibuea, 2023). Penelitian ini mengelompokkan daerah rawan stunting berdasarkan jumlah kasus stunting di Kabupaten Asahan. Hasil penelitian menunjukkan bahwa hasil jumlah *cluster* terbaik adalah dua *cluster*, yang ditentukan dengan menggunakan metode *elbow*. Evaluasi kualitas *cluster* menghasilkan nilai Davies-Bouldin Index (DBI) sebesar 0,51290 dan skor *silhouette* sebesar 0,71432, yang menunjukkan bahwa kualitas *cluster* baik dan data telah diklasifikasikan dengan tepat. Berdasarkan referensi dari penelitian terdahulu terkait *clustering*, maka penelitian ini akan menggunakan algoritma K-Means untuk mengelompokkan daerah rawan stunting di Indonesia.

K-Means adalah salah satu metode *clustering* non-hirarki untuk mengelompokkan data ke dalam satu atau lebih *cluster*, dimana data yang memiliki karakteristik yang sama dikelompokkan dalam *cluster* yang sama, dan data yang memiliki karakteristik yang berbeda dikelompokkan ke dalam *cluster* lain (Nur Afidah & Masrukan, 2023). Kelebihan dari algoritma K-Means yaitu kesederhanaan, relatif cepat dalam proses komputasi, mudah untuk adaptasi, dan mudah diimplementasikan dibandingkan dengan algoritma *clustering* lainnya (Cinderatama et al., 2022).

Dari latar belakang yang telah diuraikan tersebut, peneliti merasa perlu untuk melakukan pengelompokan daerah rawan stunting di Indonesia berdasarkan faktor-faktor penyebab stunting dengan menggunakan K-Means. Oleh karena itu, peneliti mengusulkan penelitian berjudul “**Clustering Daerah Risiko Rawan Stunting di Indonesia Menggunakan K-Means**”.

1.2. Rumusan Masalah

Berdasarkan latar belakang diatas, rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana penerapan algoritma K-Means dalam mengelompokkan daerah risiko rawan stunting di Indonesia berdasarkan faktor-faktor penyebab stunting?
2. Bagaimana menentukan jumlah *cluster* yang optimal dalam pengelompokkan daerah risiko rawan stunting di Indonesia berdasarkan faktor-faktor penyebab stunting?

1.3. Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Dataset yang digunakan bersumber dari website resmi Direktorat Jendral Bina Pembangunan Daerah (Bangda) Kementerian Dalam Negeri Republik Indonesia (Kemendagri) melalui *dashboard* cakupan layanan dan sebaran kasus stunting. *Dataset* yang digunakan meliputi variabel cakupan layanan esensial dalam penurunan stunting dan kasus stunting pada tahun 2022-2024.
2. Fokus penelitian ini adalah untuk melakukan metode *clustering* menggunakan algoritma K-Means untuk mengelompokkan daerah risiko rawan stunting di Indonesia berdasarkan variabel-variabel faktor penyebab stunting.

1.4. Tujuan Penelitian

Berdasarkan rumusan masalah diatas, maka tujuan dari penelitian ini adalah:

1. Untuk menerapkan algoritma K-Means dalam mengelompokkan daerah risiko rawan stunting di Indonesia berdasarkan faktor-faktor penyebab stunting.
2. Untuk menentukan jumlah *cluster* yang optimal dalam pengelompokkan daerah risiko rawan stunting di Indonesia berdasarkan faktor-faktor penyebab stunting.

1.5. Manfaat Penelitian

Adapun hasil dari penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Secara teoretis, penelitian ini diharapkan dapat menjadi referensi dan tambahan informasi bagi pembaca mengenai penerapan algoritma K-Means dalam *clustering* daerah risiko rawan stunting di Indonesia berdasarkan faktor-faktor penyebab stunting.

2. Secara praktis, penelitian ini dapat dijadikan referensi dan memberikan informasi yang bermanfaat bagi pemerintah Indonesia dalam pertimbangan pengambilan keputusan yang tepat. Hasil penelitian ini diharapkan dapat menjadi referensi untuk membantu pemerintah dalam mengidentifikasi daerah-daerah yang perlu diprioritaskan terlebih dahulu dalam penanganan kasus stunting di Indonesia.