BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Provinsi Jambi terletak di wilayah tengah Pulau Sumatera, dengan batas geografis ditentukan oleh garis lintang 0°45' - 2°45' LS dan garis bujur 101°10' - 104°55' BT. Dengan luas wilayah 53.435,72 kilometer persegi, Provinsi Jambi meliputi dua kota dan sembilan kabupaten, dengan kondisi geografi yang bervariasi mulai dari dataran rendah hingga pegunungan [1]. Perekonomian pada provinsi Jambi didominasi oleh sektor pertanian, perkebunan, kehutanan, dan pertambangan. Meskipun memiliki sumber daya alam yang melimpah, Provinsi Jambi masih menghadapi berbagai tantangan, termasuk masalah kesehatan masyarakat seperti stunting, yang berdampak pada perkembangan generasi mendatang.

Stunting adalah gangguan medis yang ditandai dengan pertumbuhan dan perkembangan yang tidak proporsional pada anak dibandingkan dengan usia normal mereka [2]. Kondisi stunting tidak hanya berdampak pada kesehatan tetapi juga dampak sosial dan ekonomi. Anak-anak yang menderita stunting mengalami keterlambatan perkembangan fisik dan mental, sistem kekebalan tubuh yang terganggu, masalah gizi serta kesehatan, dapat mengakibatkan turunnya prestasi akademik dan dampak jangka panjang pada produktivitas dan ekonomi[3]. Menurut WHO Pada tahun 2022, prevalensi global stunting di kalangan anak di bawah usia 5 tahun diproyeksikan mencapai 148,1 juta, sementara 41 juta anak akan mengalami

wasting dan 37 juta akan mengalami kelebihan berat badan [4]. Berdasarkan hasil Survei Status Gizi Indonesia (SGI) 2022, angka stunting di Indonesia masih cukup tinggi, yakni sebesar 21,6%. Meski mengalami penurunan dari tahun sebelumnya, yakni sebesar 24,4% pada 2021, upaya signifikan masih diperlukan untuk mencapai target penurunan stunting sebesar 14% pada 2024 [5]. Provinsi Jambi dalam hal ini juga berhasil menurunkan angka prevalensi stunting sebesar 4,4% pada Tahun 2022, sehingga menjadi 18,0%. Kondisi prevalensi stunting di Provinsi Jambi tertinggi berada di Kabupaten Batanghari sebesar 26,2% dan terendah di Kabupaten Tanjung Jabung Barat sebesar 9,9%, begitupun dengan 9 kab/kota lainnya mengalami penurunan, akan tetapi terdapat 2 Kabupaten/Kota yaitu Kabupaten Batanghari dan Kota Sungai Penuh mengalami kenaikan, dalam hal ini pemerintah terus berupaya untuk melakukan percepatan penurunan stunting dengan target yang harus dicapai yaitu 14% di tahun 2024 [6]. Pengumpulan dan analisis data menjadi kunci dalam merancang intervensi yang efektif untuk mendukung program penurunan angka stunting tersebut pada provinsi Jambi.



Sumber Data: Buku Saku SSGI 2021 dan 2022, RPJMD Provinsi Jambi 2021-2026

Gambar 1.1 Prevalensi dan Target Penurunan Stunting di Provinsi Jambi

Gambar diatas merupakan target pemerintah provinsi Jambi terhadap penurunan prevalensi stunting di 11 kabupaten kota hasil tersebut diperoleh berdasarkan musyawarah yang disampaikan melalui surat sekretaris daerah provinsi Jambi Nomor S-050/1684/BAPPEDA-3.2/VI/2022.

Berbagai upaya penanganan stunting telah dilakukan pemerintah baik di level pemerintah pusat maupun pemerintah daerah, namun upaya ini belum memperoleh hasil yang maksimal untuk menurunkan angka kejadian stunting secara signifikan [7]. Pengumpulan dan analisis data menjadi kunci dalam merancang intervensi yang efektif untuk mendukung program penurunan angka stunting di Provinsi Jambi. Proses pendataan dan pengolahan data yang dilakukan masih menggunakan pendekatan statistika dan *Microsoft excel*, belum adanya informasi mengenai pengelompokan wilayah stunting, serta banyaknya data menyebabkan lamanya proses pengolahan data [8]. Sebagai solusi untuk mempercepat pengolahan data dan memperdalam analisis data, pendekatan *Data Mining* dapat diterapkan dalam pengolahan data stunting [9].

Data Mining mengacu pada ekstraksi sistematis pengetahuan atau informasi berharga dari kumpulan data yang luas dan rumit. Tujuan utama penambangan data adalah untuk mendeteksi pola, korelasi, atau informasi yang mungkin tidak langsung terlihat dalam data, sehingga menawarkan wawasan yang lebih mendalam dan penting [10]. *Data Mining* memungkinkan pengolahan data yang besar dan kompleks untuk menemukan pola-pola tersembunyi yang dapat digunakan sebagai dasar pengambilan keputusan [11]. Ada beberapa Teknik pada *Data Mining* antara lain seperti Klasifikasi, Analisis Klaster, Analisis Korelasi, Induksi Pohon

Keputusan, dan Analisis Regresi yang dapat digunakan untuk mengidentifikasi pola dalam rangka penemuan pengetahuan dalam sekumpulan data besar [12]. Salah satu metode yang banyak digunakan dalam *Data Mining* adalah *clustering*, yang bertujuan untuk mengelompokkan data berdasarkan karakteristik yang serupa [13].

Metode *Single linkage* adalah salah satu metode *clustering* atau pengelompokan data yang menggabungkan dua data yang memiliki jarak terkecil di antara mereka, sehingga membentuk satu klaster baru [14]. *Single linkage* merupakan teknik komputasi yang digunakan dalam pengelompokan hirarki untuk mengkategorikan dokumen atau objek berdasarkan jarak terdekatnya[15]. Fitur yang konsisten dan rumit menjadi ciri algoritma *Hierarchical Clustering* (HCC) *Single linkage* untuk mengidentifikasi titik pusat klaster (*centroid*) [16]. Dengan menerapkan *Single linkage* untuk menetapkan struktur klaster awal sebelum menerapkan *K-means* dapat membantu mengurangi sensitivitas *K-means* terhadap penempatan centroid klaster awal. Pendekatan ini dapat menghasilkan hasil klasterisasi yang lebih stabil dan kinerja keseluruhan yang lebih baik [17].

K-means Clustering adalah salah satu metode *clustering* yang paling populer dan banyak digunakan dalam berbagai permasalahan, algoritma ini bekerja dengan menentukan terlebih dahulu jumlah klaster (K) dan kemudian secara iteratif memindahkan titik tengah (*centroid*) setiap klaster untuk meminimalkan varians dalam klister [10] *K-means Clustering* memiliki kelemahan yaitu sensitivitas terhadap pemilihan *centroid* awal, yang dapat menyebabkan hasil yang tidak konsisten [18]. Hasil *cluster* yang terbentuk oleh metode *K-means* ini sangat bergantung pada jumlah k yang digunakan serta inisiasi nilai pusat awal *cluster*

yang diberikan dimana dalam hal ini dapat menyebabkan masalah lokal optimal. Untuk mengatasi masalah penentuan jumlah k pada *K-means* dapat menggunakan metode *Elbow*. Metode *elbow* merupakan metode yang digunakan untuk melihat informasi dalam penentuan jumlah *cluster* terbaik dengan cara melihat persentase hasil perbandingan antara jumlah *cluster* yang membentuk siku pada suatu titik [19]. Untuk meningkatkan kualitas *cluster* yang dihasilkan dapat dilakukan penggabungan antara metode *K-means Clustering* dengan *Hierarchical clustering* [20]. Penggabungan metode *partitioning* dan *hierarchical* ini disebut metode Hierarchical *K-means* gabungan dari metode. Pendekatan *hierarchical* digunakan untuk menentukan pusat klaster awal, sedangkan *partitioning* digunakan untuk menemukan klaster ideal [21]. Metode hirarki yang digunakan pada penelitian ini adalah metode *Single linkage* yang akan di kombinasikan Bersama dengan metode *K-means.*

Penelitian analisis *cluster* pada kasus stunting telah dilakukan Pada penelitian sebelumnya yang berjudul "*Implementation of Data Classification Using K-means Algorithm in Clustering Stunting Cases*" menggunakan beberapa atribut seperti usia, berat, dan tinggi badan. Dari hasil penelitian ini, dilakukan pengelompokan status anak balita berdasarkan kategori *normal* dan *stunting*. Penelitian ini dilakukan di Desa Tegalwangi dengan menggunakan 395 dataset. Algoritma *K-means* berhasil mengelompokkan balita menjadi dua kelompok, yaitu kelompok dengan perkembangan normal dan kelompok dengan kasus *stunting*. Dari hasil evaluasi menggunakan *Davies-Bouldin Index* (DBI), diperoleh nilai

optimal sebesar 0.007 yang mendekati nol, menunjukkan bahwa pengelompokan yang dilakukan berkualitas [22].

Selanjutnya penelitian sebelumnya yang membahas perbandingan antara metode *hierarchical clustering* untuk menemukan metode dengan performa terbaik telah dilakukan, dengan judul " *K-Centroid Link: a novel hierarchical clustering linkage method* ", penelitian ini membandingkan metode *Single linkage*, *complete linkage, average linkage, centroid link* dan metode *ward* untuk menentukan pusat *cluster* yang hasilnya metode *Single linkage* menghasilkan akurasi terbaik dibandingkan metode yang lainnya. Dari metode yang di usulkan diuji dengan 24 kumpulan data acuan yang digunakan untuk pengelompokkan. Dari hasil yang diujikan sebanyak 12 dataset memberikan performa akurasi paling baik menggunakan metode *Single linkage* [23].

Selanjutnya Penelitian yang membahas penerapan metode kombinasi dari metode *Single linkage* dan *K-means clustering* telah dilakukan pada penelitian sebelumnya yang berjudul "Penerapan Metode KnA (Kombinasi *K-means* dan *Agglomerative Hierarchical Clustering*) dengan Pendekatan *Single linkage* untuk Menentukan Status Gizi pada Balita", Penelitian ini menggunakan atribut seperti jenis kelamin, usia, berat badan, dan tinggi badan dalam proses pengelompokan. Dari hasil penelitian ini, status gizi balita dikelompokkan ke dalam beberapa kategori, yaitu berat badan per umur (BB/U), tinggi badan per umur (TB/U), dan berat badan per tinggi badan (BB/TB), yang masing-masing terdiri dari empat kluster. Penelitian ini dilakukan dengan menggunakan data sebanyak 7.336 balita yang diambil dari Puskesmas Pulung Ponorogo. Metode KnA yang

menggabungkan keunggulan algoritma *K-means* dan *Agglomerative Hierarchical Clustering* berhasil mengelompokkan balita ke dalam kluster sesuai status gizinya, dengan hasil akurasi sebesar 94% untuk BB/U, 98% untuk TB/U, dan 93% untuk BB/TB. Evaluasi kualitas *clustering* dilakukan menggunakan akurasi pengelompokan berdasarkan data learning, menunjukkan bahwa metode KnA mampu menangani data dalam jumlah besar dan mengelompokkan status gizi balita secara efektif [24].

Berdasarkan penelitian sebelumnya, penggunaan kombinasi metode hirarki dan *K-means* telah terbukti menghasilkan pengelompokan yang lebih optimal. Oleh karena itu, penelitian ini akan membahas penerapan kombinasi *Single linkage* dan *K-means* dalam pengelompokan anak stunting di Provinsi Jambi. Melalui studi kasus ini, diharapkan dapat diperoleh hasil pengelompokan yang lebih stabil dan optimal untuk mengidentifikasi daerah rawan stunting di provinsi Jambi. Berdasarkan hal tersebut maka akan dilakukan penelitian dengan judul **"*DATA MINING* UNTUK PENGELOMPOKAN ANAK STUNTING DI PROVINSI JAMBI DENGAN KOMBINASI *SINGLE LINKAGE* DAN *K-MEANS CLUSTERING*".**

1.2 PERUMUSAN MASALAH

Berdasarkan latar belakang yang dipaparkan, perumusan masalah dalam penelitian ini sebagai berikut:

1. Bagaimana hasil pengelompokan data anak stunting yang telah dilakukan menggunakan metode *Single linkage* dengan *K-means Clustering* di provinsi Jambi?

2. Bagaimana kualitas *cluster* yang dihasilkan dari kombinasi metode yang dilakukan untuk pengelompokkan wilayah rentan stunting pada provinsi Jambi?

1.3 BATASAN MASALAH

Batasan masalah dibuat agar penelitian yang dilakukan tidak menyimpang terlalu jauh dari topik utama, dalam penelitian ini Batasan masalah adalah sebagai berikut:

1. Data pada penelitian merupakan data sekunder yang diperoleh dari Dinas Kesehatan Provinsi Jambi dengan jumlah data yang digunakan sebanyak 2.797 data anak stunting yang berada di provinsi jambi.
2. *Tools* yang digunakan dalam melakukan pemrosesan data untuk penerapan kombinasi algoritma *K-means* dan *Single linkage* adalah Anaconda dengan bahasa pemrograman python.
3. Atribut yang digunakan yaitu Kab/Kota, jenis kelamin, berat badan lahir, tinggi badan lahir, usia saat pengukuran, berat badan saat pengukuran, tinggi badan saat pengukuran dan nilai Z score untuk atribut TB/U yang diperoleh dari data sekunder dari Dinas Kesehatan Provinsi Jambi
4. Jumlah *cluster*/nilai K yang akan digunakan dalam penelitian ini mengikuti hasil yang didapatkan dari penerapan metode Elbow

1.4 TUJUAN PENELITIAN

Tujuan penelitian yang akan dilakukan penulis berdasarkan perumusan masalah di atas adalah sebagai berikut:

1. Menyediakan informasi mengenai *cluster* wilayah daerah rawan stunting pada provinsi Jambi.
2. Mengusulkan kombinasi metode *Single linkage* dan *K-means* untuk mengatasi kelemahan *K-means* tradisional dalam menentukan pusat klaster awal, serta mengevaluasi kualitas klaster yang dihasilkan menggunakan *Silhouette Coefficient* dan *Davies-Bouldin Indeks*.

1.5 MANFAAT PENELITIAN

Adapun manfaat yang dapat diperoleh dari penelitian yang dilakukan diantaranya adalah sebagai berikut:

1. Memberikan informasi klasifikasi wilayah rawan, sedang, dan tidak rawan stunting di Provinsi Jambi yang dapat dijadikan acuan bagi pemangku kebijakan dalam menetapkan prioritas penanganan stunting.
2. Menyajikan informasi terkait hasil kombinasi *Single linkage* dan *K-means* yang bertujuan mengatasi kelemahan metode *K-means* tradisional dalam penentuan pusat klaster awal.
3. Memberikan evaluasi dari kualitas hasil klasterisasi dengan mengukur *silhouette coefficient* dan *Davies-Bouldin Index* yang diharapkan dapat menjadi acuan bagi validasi hasil klasterisasi serta pengembangan kajian selanjutnya di bidang analisis data kesehatan masyarakat.

1.6 SISTEMATIKA PENULISAN

Penelitian ini terdiri dari 5 (lima) bab secara garis besar sistematika penulisan akan dijelaskan sebagai berikut:

BAB I : PENDAHULUAN

Bab ini membahas mengenai latar belakang, perumusan masalah, batasan masalah, tujuan dan manfaat penelitian serta sistematika penulisan.

BAB II : LANDASAN TEORI DAN TINJAUAN PUSTAKA

Di bab ini, penulis akan membahas dasar-dasar teori dan tinjauan pustaka yang digunakan sebagai referensi dalam tesis ini, yang diambil dari berbagai buku dan jurnal. Teori-teori dasar yang akan diuraikan meliputi Stunting, *Data Mining*, *Clustering*, Metode *Single linkage*, Metode *K-means*, Metode *Hierarchical K-means*. Bagian tinjauan pustaka mencakup penelitian-penelitian serupa yang dijadikan bahan perbandingan dengan penelitian yang dilakukan oleh penulis.

BAB III : METODE PENELITIAN

Bab ini akan menguraikan proses penelitian, pemilihan material, dan metodologi yang akan digunakan dalam penelitian.

BAB IV : HASIL PENELITIAN DAN PEMBAHASAN

Bab ini akan membahas secara menyeluruh analisis, perhitungan, serta hasil dari penelitian terkait pengelompokan anak stunting di Provinsi Jambi. Penelitian ini dilakukan melalui tahapan seleksi

data, pemrosesan data, dan transformasi data. Selanjutnya, data yang telah diolah akan dianalisis menggunakan dua metode *clustering* yang populer dalam *Data Mining*, yaitu *Single linkage* dan *K-means clustering*. Penelitian ini bertujuan untuk mengkombinasikan 2 metode tersebut untuk mendapatkan hasil yang lebih efektif dalam mengidentifikasi pola data untuk pengelompokan wilayah anak stunting di provinsi Jambi.

BAB V : PENUTUP

Pada bab ini menawarkan rekomendasi berdasarkan temuan penelitian serta kesimpulan yang membahas keseluruhan dari masalah dan hasil dari penelitian.