

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Gizi merupakan zat makanan yang dibutuhkan oleh tubuh dalam proses pertumbuhan dan kesehatan sesuai dengan kebutuhan tubuh (dr. Sentot Handoko, 2020). Tercatat oleh kementerian kesehatan pada tahun 2021 sampai 2022 penurunan *stunting* di Indonesia sebesar 2,8%, dan target yang diinginkan penurunan sebesar 2,7% disetiap tahunnya. Dengan demikian upaya untuk menurunkan angka *stunting* bisa mencapai 14% pada tahun 2024 nantinya sesuai target RPJMN (Admin, 2022). Menurut survei gizi di Indonesia Jawa Timur memiliki prevalensi *stunting* sebesar 19,2% pada tahun 2022, sedangkan pada tahun sebelumnya 2021 mencapai 23,5% hal ini membuktikan bahwa Jawa Timur mengurangi tingkat prevalensi sebesar 4,3 poin (Annur, 2023). Gizi menjadi peranan paling penting untuk tumbuh dan kembang bagi anak. Usia setelah kelahiran sampai usia 2 tahun merupakan periode kritis pertumbuhan pada anak yang disebut dengan *Golden Priode*. Karena pada masa itu pertumbuhan anak mulai dari fisik, kecerdasan, emosi, bahasa dan kepribadian (Makarim, 2023). Pertumbuhan di masa itu terjadi lebih pesat dibandingkan saat usia anak balita sampai memasuki usia prasekolah. Orang tua memiliki peran yang sangat penting dalam hal memberikan makanan yang memiliki nutrisi yang cukup kepada anak-anaknya. Status gizi balita harus benar-benar diperhatikan oleh orang tua dan juga pemerintah (Larasati, 2020). Wilayah Jawa Timur termasuk kedalam salah satu provinsi yang besar di Indonesia, dan datanya relatif mudah diakses dimana saja. Dalam penelitian ini difokuskan untuk mengelola data di provinsi Jawa Timur untuk dikelompokkan sesuai variabel dan data. Oleh karena itu perlu adanya *clustering* gizi balita untuk mengetahui keadaan gizi balita disetiap daerah di Jawa Timur. Pengelompokkan menggunakan sebuah algoritma *k-medoids* di yakini mampu menghitung *dataset* daerah gizi balita di Jawa Timur dengan efisien dari pada menggunakan algoritma lainnya.

Clustering merupakan salah satu metode data *mining* yang bertujuan untuk mengelompokkan suatu data kedalam kelompok/*cluster* yang memiliki karakteristik yang sama. *Clustering* tidak memerlukan data *training* untuk mengelompokkan suatu objek maka *clustering* termasuk *unsupervised machine learning*. *Clustering* sangat berbeda dengan klasifikasi, adalah tidak adanya label/variabel target dalam melakukan pengelompokan data (Edy Irwansyah, S.T., n.d.).

Algoritma yang sering digunakan dalam proses *clustering* ada beberapa yaitu *K-means*, *K-Medoids*, *Fuzzy C-Means*, *K-NN*. Algoritma *K-Medoids* merupakan algoritma yang mirip dengan *K-Means* karena sebuah algoritma yang digunakan untuk memecah sebuah *dataset* menjadi beberapa kelompok dengan beberapa kemiripan. Perbedaan antara algoritma *K-Medoids* dan *K-Means* adalah penentuan pusat *cluster*, *K-Means* menggunakan nilai rata-rata (*mean*) dari setiap *cluster* sebagai pusat *cluster* sedangkan algoritma *K-Medoids* menggunakan objek data sebagai perwakilan pusat *cluster* (Wira et al., 2019). Menurut Yanti Puspita Sari, dkk pada penelitiannya tentang pengelompokan penyebaran tuberkulosis di Kabupaten Karawang menggunakan algoritma *k-means* dengan evaluasi menggunakan *Silhouette Coefficient* menghasilkan nilai yang cukup baik dalam melakukan *clustering* (Y. P. Sari et al., 2020). Menurut Emir Luthfi dan Arie Wahyu Wijayanto dalam penelitiannya terkait perhitungan menggunakan metode *hierarchical*, *k-means* dan *k-medoids* dalam *clustering* IPM perhitungan menggunakan *k-medoids* memiliki nilai akurasi lebih baik daripada menggunakan algoritma lainnya karena memiliki hasil rasio S_w/S_b yang lebih kecil (Luthfi & Wijayanto, 2021). Banyak perhitungan *clustering* telah dilakukan di penelitian terdahulu namun yang sering digunakan yaitu algoritma *K-Means* dan *K-medoids* karena kedua algoritma tersebut mudah diterapkan dibandingkan menggunakan algoritma *Fuzzy C-Means* dan *K-NN*. Berdasarkan referensi dari penggunaan algoritma pada penelitian terdahulu dalam melakukan *clustering* maka peneliti melakukan penelitian ini menggunakan algoritma *K-medoids* untuk mengetahui kualitas dari algoritma tersebut.

Silhouette Coefficient merupakan metode yang sering digunakan untuk mengevaluasi hasil dari jumlah *cluster* optimal dari hasil perhitungan *data mining*

(Ridhwan, 2023). Nilai yang dihasilkan dari metode *Silhouette Coefficient* adalah -1 hingga 1, semakin mendekati 1 maka jumlah *cluster* dari struktur *clustering* semakin tepat namun jika mendekati -1 maka jumlah *cluster* dari struktur *clustering* yang dihasilkan *overlapping*. Menurut Allbila Rahajeng Lashiyanti, dkk pada penelitiannya terkait menggunakan *Silhouette Coefficient* dan *Elbow* dalam optimasi *clustering k-means* hasil penelitian menyatakan bahwa dengan menggunakan *Silhouette Coefficient* dalam optimasi *cluster* baik untuk mengevaluasi dari proses perhitungan (Lashiyanti et al., 2023). Sedangkan menurut Taufiq Akbar, dkk terikat penggunaan *silhouette coefficient* untuk pengujian dari hasil *clustering K-medoids* dengan menggunakan *Silhouette Coefficient* sangat membantu dalam menentukan kualitas terbaik *cluster K-Medoids* (Akbar et al., 2023).

Pengelompokan daerah *stunting* gizi balita di Jawa Timur yang memiliki beberapa variabel pendukung *dataset* dari sebuah *website* resmi Open Data Jatim. Metode *clustering* yang diterapkan dalam penelitian ini adalah algoritma *K-Medoids* karena algoritma tersebut memiliki keakuratan yang tinggi dalam proses *clustering* sehingga dengan penggunaan algoritma ini akan menjadi lebih optimal dan efisien.

Berdasarkan latar belakang tersebut maka peneliti mengusulkan penggunaan Algoritma *K-medoids* dengan metode *Silhouette Coefficient* dalam proses perhitungan agar mendapatkan akurasi yang tinggi dalam pengelompokan daerah rawan *stunting* gizi balita di Jawa Timur.

1.2 Rumusan Masalah

Berdasarkan penjelasan latar belakang diatas tentang gizi balita maka terdapat sebuah rumusan masalah yaitu bagaimana melakukan perhitungan menggunakan algoritma *K-medoids* dengan metode *Silhouette Coefficient* untuk *clustering* daerah rawan *stunting* di Jawa Timur?

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk mengetahui keefektifan dan keakuratan perhitungan *clustering* menggunakan Algoritma *K-medoids* dengan *Silhouette*

Coefficient untuk mengevaluasi kualitas perhitungan dalam *clustering* daerah rawan stunting di Jawa Timur.

1.4 Batasan Masalah

Untuk memfokuskan penelitian maka dibutuhkan batasan permasalahan dalam penelitian ini. Berikut ini batasan masalah dalam penelitian ini:

1. Pengolahan data gizi balita di Jawa Timur pada tahun 2022
2. Perhitungan *dataset* gizi balita menggunakan algoritma *K-medoids* dengan metode *Silhouette Coefficient*.

1.5 Manfaat Penelitian

- a. Manfaat praktis, sebagai informasi serta pengetahuan kepada Dinas Kesehatan dan masyarakat mengenai *clustering*/pengelompokan status gizi balita sehingga masyarakat paham dan dapat segera mengatasi permasalahan gizi untuk tumbuh dan kembang pada anak.
- b. Manfaat teoritis, sebagai pengembangan ilmu pengetahuan tentang keakuratan perhitungan data *mining* menggunakan algoritma *K-medoids* dengan metode *Silhouette Coefficient* dalam *clustering* daerah rawan *stunting* di Jawa Timur.

UNUGIRI