

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perbandingan algoritma pada klasifikasi status gizi menjadi penting karena setiap algoritma memiliki kelebihan dan kelemahan tertentu. Beberapa faktor yang perlu dipertimbangkan dalam memilih algoritma yang tepat adalah akurasi prediksi, kecepatan komputasi, kebutuhan akan data yang lengkap, skalabilitas, dan kemudahan implementasi. Selain itu, perbandingan algoritma juga diperlukan untuk memperoleh pemahaman yang lebih baik tentang kinerja algoritma pada dataset yang berbeda. Setiap dataset dapat memiliki karakteristik unik, seperti jumlah sampel, variasi fitur, atau distribusi kelas, yang dapat mempengaruhi kinerja algoritma secara keseluruhan. Selain itu, perbandingan algoritma pada klasifikasi status gizi juga dapat membantu dalam pengembangan dan peningkatan metode yang ada. Dengan membandingkan kinerja berbagai algoritma, kita dapat mengidentifikasi kelemahan dan kelebihan masing-masing algoritma. Informasi ini dapat digunakan untuk mengoptimalkan algoritma yang ada atau mengembangkan pendekatan baru yang lebih efektif.

Masalah gizi di Indonesia masih merupakan masalah nasional, yaitu kelompok usia yang sering terkena masalah gizi adalah usia balita yaitu umur 1-5 tahun. Anak yang mengalami masalah gizi pada usia dini akan mengalami gangguan tumbuh kembang, mengalami kesakitan bahkan kematian. Oleh karena itu untuk mendapatkan kesehatan yang baik, salah satunya dengan cara mengkonsumsi gizi yang cukup sesuai dengan kebutuhan tubuh, selain itu untuk mengetahui tingkat kesehatan seseorang dapat dilihat melalui status gizinya (Saleh, Faisal and Musa, 2019). Masa usia di bawah lima tahun merupakan periode yang penting bagi perkembangan dan pertumbuhan seorang anak. Usia ini mempunyai pengaruh yang cukup besar menuju tahapan selanjutnya bagi perkembangan setiap individu.

Menurut UNICEF (*United Nation of Children and Education Federation*) Masalah gizi dalam pembangunan kependudukan masih merupakan persoalan yang dianggap menjadi masalah utama dalam tatanan kependudukan dunia. Oleh

karena itu, persoalan ini menjadi salah satu butir penting yang menjadi kesepakatan global dalam MDGs (*Milleneum Development Goals*). Setiap negara secara bertahap harus mampu mengurangi jumlah balita bergizi buruk atau gizi kurang mencapai 15,5 persen setiap tahunnya (Marini *et al.*, 2022). Sebagai salah satu negara dengan kompleksitas kependudukan yang sangat beraneka ragam, Indonesia dihadapi oleh dinamika persoalan gizi buruk. Berdasarkan hasil Riset Kesehatan Dasar (Riskesdas) Kementerian Kesehatan 2018 menunjukkan 17,7% bayi usia di bawah 5 tahun (balita) masih mengalami masalah gizi. Angka tersebut terdiri atas balita yang mengalami gizi buruk sebesar 3,9% dan yang menderita gizi kurang sebesar 13,8%. Di banding hasil Riskesdas 2013, bayi yang mengalami masalah gizi turun.

Klasifikasi adalah salah satu pembelajaran yang paling umum di dalam data mining. Klasifikasi dapat didefinisikan sebagai bentuk dari analisis data yang digunakan untuk mengekstrak model yang akan digunakan untuk memprediksi label kelas. Kelas yang terdapat dalam klasifikasi merupakan atribut dalam satu set data yang paling unik yang merupakan variabel bebas yang terdapat dalam statistik. Klasifikasi data terdiri dari dua proses yaitu tahap pembelajaran dan tahap pengklasifikasian. Tahap Pembelajaran merupakan tahapan dalam pembentukan model klasifikasi, sedangkan tahap pengklasifikasian merupakan tahapan dalam penggunaan model klasifikasi yang digunakan untuk memprediksi label kelas pada data (Moch. Rizky Yuliansyah, B and Franz, 2022)

Penelitian sebelumnya yang menjadi landasan penelitian ini untuk mengklasifikasikan status gizi balita yaitu penelitian yang dilakukan oleh (Lasarudin, Gani and Tomayahu, 2022) dengan judul “Perbandingan Metode *Naïve Bayes* dan C4.5 Klasifikasi Status Gizi Bayi Balita” Hasil akurasi yang didapatkan pada penelitian ini diperoleh nilai akurasi metode *Naïve Bayes* sebesar 94.20% dan metode C4.5 sebesar 85,80% dengan menggunakan 90% data training dan 10% data testing. Selanjutnya penelitian yang dilakukan oleh (Setiawan and Triayudi, 2022) dengan judul “Klasifikasi Status Gizi Balita Menggunakan *Naïve Bayes* dan *K-Nearest Neighbor* Berbasis Web” yang menghasilkan tingkat akurasi sebesar 80.6% dengan menggunakan metode *Naïve Bayes*. Sedangkan dengan menggunakan metode K-NN mendapatkan akurasi sebesar 91.7%.

Dalam penelitian ini, penulis melakukan penelitian dengan membandingkan tiga metode untuk melakukan klasifikasi status gizi balita yaitu menggunakan metode *K-Nearest Neighbor*(K-NN), metode *Decision Tree C4.5* dan *Support Vector Machine* (SVM) menggunakan data yang diambil dari Puskesmas Negeri Agung sebanyak 1.878 data balita. Data tersebut kemudian akan diolah dengan tujuan untuk mencari nilai *accuracy*, *precision* dan *recall* yang terbaik dari ketiga metode tersebut.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah diuraikan di atas, maka penulis dapat merumuskan masalah yang ingin dijawab dalam penelitian ini yaitu Berapakah nilai tingkat akurasi, precision dan recall-nya, dari metode *K-Nearest Neighbor*(K-NN), metode *Decision Tree C4.5* dan *Support Vector Machine* (SVM) terhadap data balita dari Puskesmas Negeri Agung?

1.3 Batasan Masalah

Berikut adalah batasan masalah dalam penelitian yang dilakukan :

1. Metode yang digunakan dalam pengklasifikasian yaitu *K-Nearest Neighbor*(K-NN), metode *Decision Tree C4.5* dan *Support Vector Machine* (SVM).
2. Menggunakan software rapidminer 10.1 untuk keakuratan hasil analisis dan microsoft excel 2019 untuk membantu pengolahan data.
3. Penelitian ini menggunakan data posyandu dari Puskesmas Negeri Agung sebanyak 1878 data.
4. Atribut yang digunakan yaitu: umur, jenis kelamin, berat badan, dan tinggi badan.
5. Pengukuran antropometri yang digunakan yaitu: Berat Badan menurut Tinggi Badan (BB/TB).
6. Batas usia balita yang digunakan adalah 0-60 bulan.
7. Atribut yang digunakan dalam klasifikasi ada 6 kelas yaitu gizi buruk, kurang, baik, beresiko gizi lebih, gizi lebih dan obesitas.
8. Pengujian dilakukan dengan menghitung *confussion matrix* dengan software rapidminer.

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk membandingkan algoritma yang paling akurat dengan mencari akurasi, precision dan recall-nya dalam penentuan klasifikasi status gizi balita. Metode yang digunakan adalah metode *K-Nearest Neighbor*(K-NN), metode *Decision Tree C4.5* dan *Support Vector Machine* (SVM).

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian yang akan dilakukan dengan perbandingan algoritma ini yaitu untuk memperoleh pemahaman yang lebih baik tentang kinerja algoritma pada dataset yang berbeda dalaman melakukan klasifikasi menggunakan metode *K-Nearest Neighbor*(K-NN), metode *Decision Tree C4.5* dan *Support Vector Machine* (SVM) khususnya untuk digunakan dalam proses klasifikasi untuk menentukan status gizi pada balita. Sehingga pada nantinya hasil dari penelitian ini dapat menjadi pembanding dalam melakukan penelitian sejenis untuk ke depannya.