

BAB 1

PENDAHULUAN

1.1. Latar Belakang Masalah

Perkembangan teknologi saat ini semakin meningkat dan memberi pengaruh yang besar hampir disetiap sektor kehidupan dan kenegaraan. Proses globalisasi yang terjadi disetiap negara di dunia saat ini juga mendukung perkembangan dan penggunaan teknologi. Kebutuhan informasi akan semakin meningkat bersamaan dengan perkembangan teknologi. Semakin banyak informasi yang dibutuhkan maka data yang dibutuhkan juga semakin banyak dan jumlahnya akan semakin besar.

Jumlah data yang terus meningkat ini memerlukan beberapa metode untuk mengolah dan mengambil kesimpulan dan informasi dari data tersebut. Beberapa metode yang digunakan untuk mengolah data yang sifatnya besar untuk menemukan pola yang terdapat didalamnya diantaranya adalah: teknik klustering, analisis diskriminan, *teorema bayes*, *decision tree*, *artificial neural networks*, *support vector machine*, regresi linear, *support vector regresi*. Setiap metode tersebut memiliki algoritma-algoritma yang digunakan untuk memproses data yang ada. Namun pada kesempatan kali ini penulis mengangkat mengenai penggunaan *Algoritma C4.5* dan *Algoritma Naïve Bayes Classification* yang merupakan algoritma dari metode *decision tree* dan *teorema bayes* untuk pengklasifikasian status gizi balita. Algoritma C4.5 merupakan algoritma yang memiliki proses *learning* dan klasifikasi yang sederhana dan cepat. Secara umum, model algoritma C4.5 mempunyai tingkat akurasi yang tinggi (*Han & Kamber, 2006*). Sedangkan *Naïve Bayes Classifier (NBC)* merupakan algoritma klasifikasi yang sangat efektif (mendapatkan hasil yang tepat) dan efisien (proses penalaran dilakukan memanfaatkan *input* yang ada dengan cara yang relatif cepat). Algoritma NBC bertujuan untuk melakukan klasifikasi data pada kelas tertentu. Unjuk kerja pengklasifikasi diukur

dengan nilai *predictive accuracy* (Zhang, 2007). Alasan kenapa penulis memilih algoritma C4.5 dan algoritma NBC untuk mengklasifikasi status gizi adalah karena kedua algoritma tersebut memiliki perhitungan yang sederhana dan cepat serta memiliki tingkat akurasi yang tinggi. Sebelumnya sudah pernah ada penelitian tentang perbandingan algoritma C4.5 dan algoritma NBC yaitu penelitian yang dilakukan oleh Saufika Fatmawati dkk pada tahun 2022 yang berjudul “*Perbandingan Algoritma C 4.5 dan Algoritma Naive Bayes Untuk Klasifikasi Pekerja Migran Indonesia*”. Penelitian tersebut mengangkat tema tentang pola penempatan PMI dengan menggunakan data penempatan PMI wilayah BP3TKI Semarang. Algoritma Naïve Bayes digunakan untuk mengklasifikasikan data PMI dengan menghitung probabilitas dari data training dan data testing. Algoritma C4.5 digunakan untuk memprediksi dengan mengolah variabel usia, gender, pendidikan, staus_perkawinan, pendidikan, negara_tujuan, status_PMI, sektor_pekerjaan. Algoritma C 4.5 mampu memprediksi lebih baik dengan tingkat akurasi sebesar 84.84% sedangkan Algoritma Naïve Bayes menghasilkan nilai akurasi sebesar 58.29%

Setiap orang memiliki status gizi yang berbeda sesuai dengan usianya. Status gizi dapat ditentukan melalui pemeriksaan laboratorium maupun secara antropometri. Antropometri merupakan cara penentuan status gizi yang paling mudah dan murah. Pengukuran antropometri adalah pengukuran yang digunakan untuk menentukan keadaan gizi seseorang. Agar memperoleh hasil yang tepat, diberikan suatu patokan sebagai pedoman. Adapun pedoman antropometri bagi penentuan keadaan gizi merupakan parameter yang dipilih dan dianjurkan, yang meliputi penilaian terhadap usia dan berat badan, panjang badan, atau tinggi badan, dan lingkaran lengan atas (DKK, 2007).

Berdasarkan masalah yang telah diuraikan, maka penulis mengangkat judul “*KOMPARASI ALGORITMA C4.5 DENGAN NAÏVE BAYES CLASSIFIER (NBC) UNTUK KLASIFIKASI STATUS GIZI*”

BALITA” agar bisa membandingkan tingkat performa dari masing-masing algoritma dalam pengklasifikasian gizi balita.

1.2. Rumusan Masalah

Dalam penelitian kali ini penulis membandingkan kinerja algoritma C4.5 dan Naive Bayes dalam memprediksi status gizi balita. Penulis mengamati apakah terdapat perbedaan signifikan dalam akurasi klasifikasi antara algoritma C4.5 dan Naive Bayes dalam konteks status gizi balita. Kemudian bagaimana tingkat presisi dan recall kelas (tingkat pengenalan kelas) dari masing-masing algoritma dalam melakukan klasifikasi status gizi balita. Meneliti apakah faktor-faktor tertentu seperti berat badan, tinggi badan, dan umur balita memiliki pengaruh yang berbeda terhadap kinerja kedua algoritma dalam klasifikasi status gizi balita.

1.3. Ruang Lingkup

1. *Output* dari program ini berupa informasi klasifikasi status gizi balita apakah status gizinya normal, gemuk, obesitas, kurus, atau sangat kurus, serta hasil pengukuran performa yang terdiri dari akurasi, presisi dan recall kelas dari algoritma C4.5 dengan algoritma NBC.
2. *Algoritma* yang digunakan yaitu algoritma C4.5 yang merupakan algoritma dari metode *Decision Tree* dan algoritma NBC yang merupakan algoritma dari *Theorema Bayes*.
3. Untuk menguji perbandingan kedua algoritma tersebut menggunakan aplikasi Rapid Miner.

1.4. Tujuan Penelitian

Tujuan yang ingin dicapai adalah mengetahui perbandingan performa algoritma C4.5 dan algoritma NBC dalam pengklasifikasian status gizi balita, yang meliputi tingkat akurasi, presisi dan recall kelas.

1.5. Manfaat Penelitian

Manfaat dari penelitian ini adalah dapat memberikan wawasan yang mendalam mengenai kinerja relatif dari algoritma C4.5 dan Naive Bayes

dalam tugas klasifikasi status gizi balita. Hasilnya dapat membantu peneliti, praktisi, atau pengembang sistem untuk memilih algoritma yang paling cocok untuk tugas tersebut.

Dengan memahami kelebihan dan kelemahan masing-masing algoritma, penelitian ini dapat menjadi dasar untuk mengoptimalkan kinerja model klasifikasi. Hal ini dapat mencakup penyesuaian parameter algoritma, pemanfaatan fitur yang lebih efektif, atau strategi lain untuk meningkatkan akurasi prediksi.

1.6. Sistematika Penulisan

Tugas akhir ini ditulis sesuai sistematika penulisan sebagai berikut:

1. BAB I : PENDAHULUAN
Bab ini berisi tentang latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat, metodologi penelitian dan sistematika penulisan.
 2. BAB II : TINJAUAN PUSTAKA DAN DASAR TEORI
Bab ini menjelaskan tentang tinjauan pustaka dan teori-teori yang terkait dengan permasalahan yang diambil.
 3. BAB III : METODOLOGI PENELITIAN
Bab ini berisi tentang metodologi penelitian yang digunakan dalam penelitian tugas akhir ini
 4. BAB IV : IMPLEMENTASI DAN PEMBAHASAN
Bab ini menjelaskan seluruh proses yang sudah dirancang pada bab sebelumnya, analisa data, pengolahan data dan perhitungan algoritma.
 5. BAB V : SIMPULAN DAN SARAN
Bab ini menjelaskan tentang hasil dari penelitian yang menjawab masalah dan tujuan penelitian serta saran yang berisi tentang alternatif pengembangan lebih lanjut.
- DAFTAR PUSTAKA
Bab ini berisi tentang daftar pustaka yang diambil.
- LAMPIRAN. Bab ini berisi tentang lampiran -lampiran.