

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pola konsumsi masyarakat Indonesia menunjukkan kecenderungan meningkatnya konsumsi makanan yang tinggi kalori, gula, dan lemak. Kondisi ini sejalan dengan hasil Riset Kesehatan Dasar yang mencatat bahwa prevalensi obesitas pada orang dewasa telah mencapai 21,8% dan terus mengalami peningkatan[1]. Fenomena tersebut mengindikasikan bahwa sebagian masyarakat masih mengalami kesulitan dalam menafsirkan dan mengevaluasi tingkat kesehatan makanan berdasarkan informasi kandungan nutrisi yang tersedia, khususnya pada makanan olahan yang memiliki komposisi gizi kompleks dan tidak mudah dipahami secara langsung.

Pola makan yang tidak seimbang berdampak langsung terhadap peningkatan kasus penyakit tidak menular (PTM), seperti diabetes melitus, hipertensi, dan penyakit jantung. Kementerian Kesehatan menyatakan bahwa tingginya konsumsi makanan tinggi gula dan lemak, disertai dengan rendahnya aktivitas fisik, merupakan faktor utama yang berkontribusi terhadap menurunnya kondisi kesehatan masyarakat[1]. Selain itu, media nasional juga melaporkan bahwa permasalahan obesitas mulai banyak ditemukan pada kelompok usia muda, terutama generasi Alpha, yang cenderung mengonsumsi makanan cepat saji dan makanan olahan secara berlebihan[2].

Berbagai penelitian di tingkat global menunjukkan bahwa pola makan yang tidak sehat merupakan salah satu faktor risiko terbesar terhadap gangguan kesehatan. Konsumsi gula berlebih, lemak jenuh, serta asupan natrium yang melebihi batas aman terbukti meningkatkan risiko penyakit metabolik dan kardiovaskular[3]. Hal ini menegaskan bahwa pemilihan makanan yang tepat berdasarkan kandungan nutrisi merupakan langkah penting dalam upaya pencegahan penyakit sejak dini.

Sebaliknya, penerapan pola konsumsi makanan sehat memberikan manfaat yang signifikan bagi kesehatan tubuh. Konsumsi makanan dengan kandungan nutrisi seimbang, seperti protein, serat, vitamin, dan mineral, berperan dalam menjaga berat badan ideal, meningkatkan sistem kekebalan tubuh, serta menurunkan risiko penyakit tidak menular[4]. Selain itu, Kementerian Kesehatan Republik Indonesia melalui pedoman gizi seimbang juga menekankan bahwa konsumsi makanan sehat berkontribusi terhadap peningkatan kualitas hidup dan produktivitas individu[5].

Kemajuan teknologi data memberikan peluang untuk mendukung upaya tersebut melalui penerapan machine learning. Sejumlah penelitian menunjukkan bahwa algoritma machine learning mampu mengolah data nutrisi dan melakukan klasifikasi makanan sehat maupun tidak sehat secara akurat[6]. Oleh karena itu, penelitian ini mengembangkan model klasifikasi makanan sehat dan tidak sehat menggunakan algoritma Random Forest dan Gradient Boosting sebagai sistem pendukung keputusan yang membantu menginterpretasikan data nutrisi makanan yang tersedia, sehingga dapat memberikan penilaian tingkat kesehatan makanan secara objektif dan terukur.

Selain itu, penelitian ini juga mengevaluasi kinerja kedua algoritma pada dataset nutrisi yang memiliki distribusi kelas tidak seimbang, serta mengkaji pengaruh penerapan teknik penyeimbangan data terhadap performa model klasifikasi. Hasil penelitian menunjukkan bahwa algoritma Random Forest tanpa penerapan resampling memberikan performa terbaik dibandingkan dengan skenario oversampling dan undersampling. Temuan ini menunjukkan bahwa pemilihan metode klasifikasi dan strategi pengolahan data perlu disesuaikan dengan karakteristik dataset nutrisi makanan yang digunakan.

1.2 Rumusan Masalah

Adapun rumusan masalah dalam penelitian ini adalah sebagai berikut.

1. Bagaimana proses pengolahan data nutrisi makanan dapat digunakan untuk membangun model klasifikasi makanan sehat dan tidak sehat?
2. Seberapa baik kinerja algoritma Random Forest dan Gradient Boosting dalam memprediksi kategori makanan sehat berdasarkan kandungan nutrisi?
3. Faktor nutrisi apa saja yang paling berpengaruh terhadap hasil prediksi kesehatan makanan berdasarkan analisis feature importance kedua model tersebut?

1.3 Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah sebagai berikut.

1. Penelitian hanya menggunakan data nutrisi makanan seperti kalori, protein, karbohidrat, lemak, serat, gula, natrium, kolesterol, serta ukuran saji sebagai dasar penentuan kategori sehat atau tidak sehat. Faktor eksternal seperti pola konsumsi, kondisi medis, atau preferensi diet individu tidak dibahas.
2. Model klasifikasi hanya menggunakan dua algoritma, yaitu *Random Forest* dan *Gradient Boosting*. Algoritma lain seperti SVM, KNN, atau Deep Learning tidak dianalisis secara mendalam pada penelitian ini.
3. Kategori sehat dan tidak sehat ditentukan berdasarkan aturan nutrisi tertentu, yang sudah diturunkan menjadi label pada dataset, sehingga penelitian tidak mengembangkan standar nutrisi baru.
4. Data yang digunakan merupakan dataset public dari platform kaggle yang telah dimodifikasi, sehingga hasil penelitian terbatas pada karakteristik dataset tersebut.

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut.

1. Menganalisis proses pengolahan data nutrisi makanan, mulai dari pembersihan data, penanganan ketidakseimbangan kelas, hingga normalisasi fitur, sehingga data siap digunakan untuk proses klasifikasi.
2. Mengembangkan dan membangun model klasifikasi makanan sehat dan tidak sehat menggunakan algoritma *Random Forest* dan *Gradient Boosting* berdasarkan fitur-fitur nutrisi seperti kalori, lemak, gula, natrium, protein, dan serat.
3. Mengevaluasi performa model yang dibangun menggunakan metrik akurasi, presisi, recall, F1-score.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut.

a. Manfaat Teoritis

1. Memberikan kontribusi ilmiah dalam pengembangan model klasifikasi nutrisi makanan berbasis machine learning, khususnya penggunaan algoritma *Random Forest* dan *Gradient Boosting* pada data gizi.
2. Menambah referensi akademik mengenai penerapan teknik penyeimbangan data seperti SMOTE dalam penelitian klasifikasi kesehatan makanan.
3. Menyediakan dasar teoritis bagi penelitian selanjutnya terkait prediksi kualitas makanan, diet, atau kesehatan menggunakan pendekatan komputasional dan data nutrisi.

b. Manfaat Praktis

1. Membantu pengguna umum dalam mengidentifikasi apakah suatu makanan termasuk kategori sehat atau tidak sehat berdasarkan kandungan nutrisi secara cepat dan mudah.

2. Menjadi dasar pengembangan sistem lebih lanjut, misalnya integrasi dengan mobile app, rekomendasi menu sehat, atau analisis pola konsumsi untuk pengguna di masa depan.

1.6 Sistematika Penulisan

Sistematika penulisan skripsi ini disusun untuk memberikan gambaran secara umum mengenai isi yang dibahas pada setiap bab. Penulisan skripsi ini terdiri dari lima bab, dengan sistematika sebagai berikut.

Bab I Pendahuluan, berisi latar belakang penelitian yang menjelaskan permasalahan terkait pola konsumsi makanan dan dampaknya terhadap kesehatan. Bab ini juga memuat rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan sebagai gambaran alur penyusunan skripsi.

Bab II Tinjauan Pustaka, berisi studi literatur yang berkaitan dengan pola makan, makanan sehat dan tidak sehat, serta penerapan machine learning dalam bidang kesehatan dan nutrisi. Selain itu, bab ini juga memuat dasar teori yang digunakan sebagai landasan penelitian, meliputi konsep data nutrisi makanan, klasifikasi, confusion matrix, algoritma Random Forest, Gradient Boosting, metode SMOTE, serta metrik evaluasi model klasifikasi.

Bab III Metode Penelitian, menjelaskan tahapan dan prosedur penelitian yang dilakukan. Bab ini mencakup objek penelitian, sumber dan karakteristik data, tahapan pra-pemrosesan data, penanganan ketidakseimbangan kelas, pembangunan model klasifikasi menggunakan Random Forest dan Gradient Boosting, serta metode evaluasi yang digunakan untuk menilai performa model.

Bab IV Hasil dan Pembahasan, menyajikan hasil dari seluruh tahapan penelitian yang telah dilakukan. Bab ini membahas hasil pra-pemrosesan data, hasil penerapan SMOTE, performa model klasifikasi berdasarkan metrik evaluasi, perbandingan kinerja antara Random Forest dan Gradient Boosting, serta analisis feature importance terhadap kandungan nutrisi makanan.

Bab V Penutup, berisi kesimpulan yang diperoleh berdasarkan hasil penelitian dan pembahasan yang telah dilakukan. Selain itu, bab ini juga memuat saran yang dapat digunakan sebagai bahan pertimbangan untuk pengembangan penelitian selanjutnya.

