

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG MASALAH

Obesitas adalah penumpukan lemak secara berlebihan yang terjadi di daerah perut. Meningkatnya populasi obesitas dikaitkan dengan perubahan pola makan dan gaya hidup tidak sehat. Adanya Perubahan perilaku pola makan yang tidak sehat, seperti pola makan yang mengkonsumsi lemak berlebih, bahan pengawet, dan rendah serat dapat memicu obesitas. *World Health Organization* melaporkan bahwa pada tahun 2022, 2,5 miliar orang dewasa berusia 18 tahun ke atas mengalami kelebihan berat badan, termasuk lebih dari 890 juta orang dewasa yang menderita obesitas. Hal ini setara dengan 43% orang dewasa berusia 18 tahun ke atas (43% pria dan 44% wanita) yang kelebihan berat badan meningkat dari tahun 1990, ketika 25% orang dewasa berusia 18 tahun ke atas mengalami kelebihan berat badan[1]. Indonesia menempati urutan kedua setelah Singapura dengan jumlah remaja obesitas terbesar yaitu 12,2% kemudian Thailand sebesar 8%, Malaysia sebesar 6% dan Vietnam sebesar 4,6%[2].

Penyebab obesitas sangat beragam, termasuk faktor biologis, Perkembangan, lingkungan, perilaku, dan genetik. Komorbiditas yang sering terjadi pada orang dewasa dengan obesitas antara lain diabetes melitus tipe 2, hipertensi, penyakit hati berlemak non-alkohol, apnea tidur obstruktif, dislipidemia, dan sindrom metabolik. Kondisi ini juga meningkat seiring dengan meningkatnya prevalensi obesitas pada anak-anak dan remaja. Selain itu, obesitas pada anak-anak

dan remaja dapat menyebabkan gangguan psikologis seperti depresi, kecemasan rendah diri, masalah sosial, dan gangguan makan. Ketidak seimbangan energi yakni asupan kalori yang lebih banyak dibandingkan kalori yang dikeluarkan merupakan penyebab utama obesitas pada anak-anak dan remaja[3]. Obesitas bukan hanya mengenai jumlah lemak yang ada di dalam tubuh tetapi juga distribusi penyimpanan lemak. Kelebihan jumlah lemak tubuh umumnya akan disimpan di jaringan adiposa di bagian bawah kulit atau rongga perut. Laki-laki dan perempuan mempunyai distribusi lemak tubuh yang berbeda-beda. Obesitas abdominal atau android (tipe buah apel) lebih sering terjadi pada laki-laki karena timbunan lemak pada bagian perut, sedangkan obesitas ginoid (tipe buah pear) lebih sering terjadi pada perempuan karena lemak tertimbun pada tubuh bagian bawah terutama pada daerah pinggul dan paha[4].

Data RISKESDAS tahun 2018 menunjukkan bahwa sebanyak 21,8% penduduk di Indonesia yang berada pada usia diatas 18 tahun menderita obesitas. Sedangkan menurut jenis kelamin prevalensi obesitas lebih tinggi pada wanita. Menurut data BPS, orang dengan usia 15 tahun ke atas yang bekerja sebagai pegawai atau pekerja kantoran di Indonesia pada tahun 2018 sebanyak 49 juta, sedangkan pada tahun 2021 bulan Februari ada sebanyak 48 juta. Berdasarkan data WHO dan RISKESDAS, jumlah penderita obesitas cukup tinggi di usia 18 tahun ke atas. Data BPS di setiap tahunnya sampai 2021 juga menunjukan bahwa jumlah orang dengan usia 18 tahun ke atas yang bekerja sebagai pekerja kantoran atau pegawai adalah yang paling banyak menderita obesitas dibandingkan dengan pekerjaan lain[5].

Perkembangan ilmu pengetahuan dan teknologi ada peluang besar untuk membangun model data mining yang dapat membantu mengidentifikasi tingkat obesitas dengan lebih efisien. Data mining merupakan sebuah teknik yang dapat dipergunakan untuk membantu penyelesaian permasalahan tersebut. Pada data mining banyak digunakan untuk melakukan proses pengolahan data. Klasifikasi adalah proses pembelajaran fungsi atau model pada sekumpulan data latih sehingga model tersebut dapat digunakan untuk memprediksi klasifikasi data uji[6].

Pada penelitian yang dilakukan oleh Jiawei dalam jurnal (Miranda & Julisar, 2018) digunakan Body Fat Extended Dataset sebagai dasar untuk melakukan klasifikasi tingkat obesitas pada pria. Algoritma *Naïve Bayes* dengan beberapa kriteria yang didapat dari pengukuran antropometri pada pria beserta perhitungan Indeks Massa Tubuh (IMT) dan digunakanlah perhitungan persen lemak tubuh berdasarkan perhitungan lingkar perut dan lingkar leher sebagai kriteria utama, dengan mengimplementasikan algoritma *Naïve Bayes* melakukan perhitungan persentase lemak tubuh sehingga bisa mengklasifikasi dan memprediksi tingkat obesitas pada pria. Kemudian hasil dari klasifikasi tersebut akan menghasilkan dua buah kategori yaitu kategori obesitas dan tidak obesitas berdasarkan variabel. Dari hasil pengujian klasifikasi pada rasio 70% training dan 30% testing menggunakan *naïve bayes* didapatkan hasil Accuracy 88%, Precision (Obesitas) 73%, sedangkan Precision (Non Obesitas) 98%, Recall (Obesitas) 96% dengan Recall (Non Obesitas) 85% dan hasil F1-Score sebesar 0.830188679. Hasil tersebut menunjukkan skala rata-rata diatas 80% sehingga dapat dikatakan bahwa algoritma *Naïve Bayes Classifier (NBC)* mampu mengklasifikasi obesitas pada pria dengan

baik. Pada penelitian ini penulis menggunakan metode klasifikasi algoritma Naïve Bayes. Menurut Jiawei dalam jurnal (Miranda & Julisar, 2018) menjelaskan bahwa keunggulan menggunakan metode naïve bayesclassifier dibandingkan metode lainnya adalah merupakan metode pengklasifikasian statistic yang dapat digunakan untuk memprediksi probabilitas keanggotaan dari suatu kelas, selain itu terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database yang besar. Ciri utama dari Naïve Bayes Classifier (NBC) adalah asumsi yang sangat kuat (naïf) akan independensi dari masing-masing kondisi atau kejadian [7].

Namun, dibandingkan dengan dataset ObesityPrediction.csv yang digunakan dalam penelitian ini, dataset Body Fat Extended memiliki beberapa keterbatasan. Body Fat Extended hanya mencakup responden pria dan berfokus pada data fisik saja, seperti pengukuran lingkaran leher, lingkaran perut, tinggi, dan berat badan. Dataset tersebut tidak mempertimbangkan faktor gaya hidup seperti pola makan, kebiasaan merokok, konsumsi air putih, maupun aktivitas fisik harian, sehingga kurang mewakili penyebab obesitas secara menyeluruh dalam populasi umum. Dalam penelitian ini, penulis menggunakan dataset ObesityPrediction.csv dari Kaggle yang lebih lengkap karena mencakup data fisik sekaligus atribut gaya hidup seperti konsumsi makanan cepat saji, jumlah makanan harian, kebiasaan olahraga, serta moda transportasi yang digunakan sehari-hari. Dataset ini juga mencakup responden pria dan wanita, sehingga lebih representatif dari sisi demografi. Kategori obesitas dalam dataset ini dapat dibagi menjadi beberapa kelas seperti Underweight, Normal, Overweight, dan Obese, yang dapat disesuaikan lebih lanjut tergantung pada kebutuhan klasifikasi. Dengan kombinasi atribut fisik dan

gaya hidup, serta cakupan responden yang lebih luas, analisis klasifikasi yang dilakukan pada dataset ini menjadi lebih kompleks, menyeluruh, dan relevan terhadap kondisi obesitas yang sebenarnya di masyarakat.

Berdasarkan uraian latar belakang diatas maka penulis akan melakukan penelitian dengan judul: “IMPLEMENTASI ALGORITMA *NAÏVE BAYES* UNTUK KLASIFIKASI TINGKAT OBESITAS”

1.2 RUMUSAN MASALAH

Berdasarkan penjelasan latar belakang diatas, maka rumusan masalah dari penelitian ini adalah :

Bagaimana implementasi algoritma *Naïve Bayes* dalam klasifikasi tingkat obesitas berdasarkan data gaya hidup dan faktor kesehatan, serta bagaimana tingkat akurasi dalam mendukung deteksi dini obesitas?

1.3 BATASAN MASALAH

Dalam penelitian ini, beberapa batasan masalah ditetapkan untuk menjaga fokus dan kejelasan analisis, sebagai berikut:

1. Penelitian ini hanya fokus pada penerapan algoritma *Naïve Bayes* dalam klasifikasi tingkat obesitas.
2. Penelitian ini akan menggunakan dataset dari Kaggle yaitu, *Obesity prediction*(<https://www.kaggle.com/datasets/ruchikakumbhar/obesity-prediction>).

3. Tidak akan dilakukan perbandingan mendalam dengan algoritma *machine learning* lainnya, kecuali sebagai referensi evaluasi model.
4. Hasil penelitian yang dipaparkan adalah *akurasi*, *presisi*, *recall*, dan *F1-score*.

1.4 TUJUAN PENELITIAN DAN MANFAAT PENELITIAN

1.4.1 TUJUAN PENELITIAN

Berdasarkan rumusan masalah yang didapat, maka tujuan penelitian ini Adalah sebagai berikut:

1. Mengimplementasikan algoritma *Naïve Bayes* untuk klasifikasi tingkat obesitas berdasarkan data gaya hidup dan faktor kesehatan.
2. Melakukan pra-pemrosesan data, termasuk pembersihan dan penambahan fitur Body Mass Index (BMI), guna meningkatkan kualitas data sebelum dilakukan pemodelan.
3. Mengevaluasi performa algoritma *Naïve Bayes* menggunakan metrik akurasi, presisi, recall, F1-score, serta 10-Fold Cross Validation untuk mengetahui tingkat keakuratan dan konsistensi model.
4. Mengembangkan aplikasi berbasis Streamlit sebagai media prediksi interaktif untuk mendukung deteksi dini obesitas.

1.4.2 MANFAAT PENELITIAN

Adapun penelitian ini memiliki manfaat sebagai berikut :

1. Menerapkan algoritma *Naïve Bayes* untuk mengklasifikasikan tingkat obesitas berdasarkan data gaya hidup dan kondisi fisik individu.
2. Penelitian ini diharapkan dapat membantu masyarakat dan tenaga kesehatan dalam mendeteksi obesitas secara dini, sehingga meningkatkan upaya pencegahan dan pengelolaan kesehatan.
3. Penelitian ini dapat memberikan kontribusi pada pengembangan metode klasifikasi menggunakan Machine Learning, khususnya *Naïve Bayes*, dalam bidang kesehatan, serta menjadi referensi untuk penelitian lebih lanjut terkait klasifikasi risiko obesitas dan penyakit terkait.

1.5 SISTEMATIKA PENELITIAN

Sistematika dibuat untuk memberi suatu gambaran yang jelas terhadap penulisan ilmiah agar menghindari terjadinya pembahasan diluar dari tema dan judul penelitian. Secara garis besar penulisan ilmiah terdiri dari 5 bab, dapat dilihat dari sistematika penulisan sebagai berikut :

BAB I PENDAHULUAN

Bab ini merupakan bab pendahuluan yang membahas tentang latar belakang masalah, perumusan masalah, batasan masalah, tujuan dan manfaat penelitian serta sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini memuat tentang landasan teori yang mendasari pembahasan laporan secara khusus dan berisikan definisi - definisi yang melandasi penelitian yang dilakukan dengan

melakukan studi leterature sebagai dasar dalam melakukan implementasi algoritma dan klasifikasi.

BAB III METODOLOGI PENELITIAN

Bab ini membahas secara rinci kerangka kerja penelitian, metode atau pendekatan yang digunakan, dan tools (alat bantu) yang digunakan dalam menjawab masalah penelitian untuk mencapai suatu tujuan.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan pembahasan hasil yang diperoleh dan model deteksi menggunakan klasifikasi yang dibangun dalam penelitian ini. Serta evaluasi dan penguji sistem yang telah dihasilkan untuk melihat kesesuaian dengan tujuan yang diharapkan.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dan saran-saran dari penelitian guna mengembangkan implementasi algoritma Naïve Bayes dalam klasifikasi tingkat obesitas yang telah di uji coba.