

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Obesitas telah menjadi salah satu tantangan kesehatan masyarakat yang paling serius di abad ke-21. *World Health Organization* (WHO) melaporkan bahwa pada tahun 2022, sebanyak 2.5 miliar orang dewasa berusia 18 tahun keatas mengalami kelebihan berat badan dan 890 juta termasuk menderita obesitas [1]. Di Indonesia, berdasarkan berbagai studi, tingkat obesitas pada orang dewasa mengalami peningkatan yang signifikan dalam beberapa dekade terakhir, mencerminkan perubahan gaya hidup dan pola makan yang semakin modern. Hal ini menandakan perlunya intervensi kesehatan untuk mengurangi risiko obesitas di masa depan [2].

Peningkatan tingkat obesitas ini memiliki dampak yang sangat signifikan terhadap kesehatan masyarakat dan sistem kesehatan [3]. Obesitas berkontribusi terhadap peningkatan risiko berbagai penyakit kronis, termasuk diabetes tipe 2, penyakit kardiovaskular, dan berbagai jenis kanker. Studi ini juga mengungkapkan bahwa obesitas memberikan beban ekonomi yang signifikan pada sistem kesehatan, dengan meningkatnya kebutuhan akan perawatan terkait penyakit kronis [4]. Hal ini menunjukkan bahwa obesitas tidak hanya menjadi masalah kesehatan individu tetapi juga berdampak pada ekonomi dan sistem kesehatan secara keseluruhan.

Gaya hidup modern telah diidentifikasi sebagai faktor utama dalam peningkatan kasus obesitas. Perubahan dalam pola makan yang kaya akan makanan berkalori tinggi, berkurangnya aktivitas fisik, serta meningkatnya ketergantungan pada teknologi digital, semuanya berkontribusi pada kenaikan Indeks Massa Tubuh (IMT) [5]. Hal ini menunjukkan bahwa lingkungan yang mendukung konsumsi makanan cepat saji dan gaya hidup yang semakin tidak aktif adalah faktor utama dalam meningkatnya jumlah kasus obesitas, menunjukkan bahwa obesitas adalah hasil dari perubahan gaya hidup yang semakin kurang aktif dan berisiko bagi kesehatan [6].

Dalam upaya mengatasi masalah obesitas, teknologi *Machine Learning* telah menunjukkan potensi yang menjanjikan. Penelitian oleh D. Sitanggang and S. Sherly [7] menggunakan algoritma *Super Vector Machine* (SVM) untuk klasifikasi obesitas mencapai akurasi 71.80% dengan menganalisis berbagai parameter faktor risiko obesitas. Sementara itu, Sibi dan Widiarti [8] mengembangkan model klasifikasi obesitas menggunakan algoritma *K-Nearest Neighbor* (KNN) yang mampu memprediksi risiko obesitas pada individu dengan akurasi 79,96% berdasarkan data gaya hidup dan kondisi fisik.

Penggunaan teknologi *Machine Learning* dalam klasifikasi obesitas memiliki beberapa keunggulan signifikan. Sementara itu, Lestari et al. [9] mengimplementasikan 2 algoritma *Machine Learning* yaitu *K-Nearest Neighbor* (KNN) dan *Naive Bayes Classifier* (NBC) untuk mengidentifikasi faktor-faktor risiko obesitas pada anak-anak disabilitas. Dalam studi ini, algoritma *K-Nearest*

Neighbor (KNN) menunjukkan hasil terbaik dengan tingkat akurasi sebesar 68% dibandingkan dengan menggunakan algoritma *Naive Bayes Classifier* (NBC) yang mendapatkan akurasi sebesar 48%. Hasil ini menunjukkan bahwa metode *Nearest Neighbor* (KNN) efektif dalam mendeteksi tingkat obesitas, memberikan akurasi yang lebih tinggi dibandingkan dengan metode lain seperti *Naive Bayes Classifier* (NBC).

Studi komprehensif oleh Naqibuzzahidin dan Fatah [10] menggunakan algoritma *Random Forest* untuk menganalisis faktor risiko obesitas pada individu. Penelitian ini menggunakan dataset berisi 2.086 entri dengan 15 fitur terkait faktor risiko obesitas seperti gaya hidup, pola makan, dan riwayat kesehatan, dan menghasilkan akurasi 79,81%. Hasil menunjukkan bahwa konsumsi makanan tinggi kalori, kurangnya aktivitas fisik, dan riwayat keluarga obesitas merupakan faktor dominan dalam meningkatkan risiko obesitas. Penelitian ini mendukung pengembangan kebijakan kesehatan untuk pencegahan obesitas dengan menekankan perubahan gaya hidup yang lebih sehat.

Kebutuhan akan sistem klasifikasi yang akurat dan efisien menjadi semakin penting mengingat kompleksitas faktor-faktor yang mempengaruhi obesitas [11]. Pendekatan *Machine Learning* dapat membantu mengidentifikasi pola-pola tersembunyi dalam data faktor risiko obesitas yang sulit dideteksi menggunakan metode statistik konvensional. Kombinasi faktor risiko obesitas seperti pola tidur, tingkat stres, dan kebiasaan makan memiliki interaksi *non-linear* yang kompleks dalam mempengaruhi tingkat obesitas [12].

Lebih lanjut, penelitian oleh Yamantri dan Ahmad [13] menggunakan algoritma C4.5 untuk mengidentifikasi faktor risiko obesitas pada penduduk dewasa di Indonesia. Menggunakan data dari 1000 responden, analisis menunjukkan bahwa pola makan yang buruk dan rendahnya tingkat aktivitas fisik merupakan faktor utama yang berkontribusi pada obesitas, sementara usia dan riwayat kesehatan keluarga juga memiliki pengaruh signifikan. Model C4.5 yang dihasilkan memiliki akurasi prediksi sebesar 85%, menunjukkan potensinya dalam mengidentifikasi individu dengan risiko tinggi obesitas.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengembangkan sistem klasifikasi tingkat obesitas menggunakan algoritma *Machine Learning* dengan mempertimbangkan berbagai faktor risiko obesitas. Pendekatan ini diharapkan dapat memberikan kontribusi signifikan dalam upaya pencegahan dan penanganan obesitas yang lebih tepat sasaran, serta memberikan pemahaman yang lebih baik tentang interaksi kompleks antara berbagai faktor risiko obesitas yang mempengaruhi suatu individu.

1.2 RUMUSAN MASALAH

Berdasarkan latar belakang diatas, pokok permasalahan yang akan diteliti adalah sebagai berikut :

1. Bagaimana mengidentifikasi faktor-faktor risiko obesitas pada individu yang berpengaruh signifikan terhadap tingkat obesitas berdasarkan data yang tersedia?

2. Bagaimana mengimplementasikan dan membandingkan kinerja algoritma *Machine Learning* dalam mengklasifikasikan faktor risiko obesitas pada individu?
3. Bagaimana efektifitas berbagai algoritma *Machine Learning* dalam memprediksi risiko obesitas, dan algoritma mana yang paling akurat dalam mengklasifikasikan obesitas berdasarkan kondisi individu?

1.3 BATASAN MASALAH

Untuk memastikan pembahasan tetap terfokus pada topik dan judul penelitian, batasan kajian ditetapkan sebagai berikut :

1. Dataset yang digunakan dalam penelitian ini berasal dari *Repository Kaggle* dengan judul *Obesity Risk Prediction* [14]. Dataset ini memiliki total 2086 data, terdiri dari 15 fitur dan 1 target dengan 7 kelas.
2. Penelitian ini menggunakan algoritma *Machine Learning* untuk membangun model klasifikasi multikelas faktor risiko obesitas pada individu. Algoritma yang akan digunakan dibatasi pada algoritma-algoritma *Machine Learning* yang umum dan telah terbukti efektif untuk tugas klasifikasi, seperti KNN, *Random Forest*, dan C4.5.
3. Penelitian ini berfokus pada mengembangkan model klasifikasi dan tidak mencakup pengembangan aplikasi atau sistem terintegrasi untuk memantau dan mengelola tingkat obesitas. Integrasi model klasifikasi ke dalam aplikasi atau sistem tersebut berada diluar lingkup penelitian ini.

4. Penelitian ini akan mengevaluasi performa model klasifikasi menggunakan matriks konfusi (*confusion matrix*), *k-fold cross validation*, dan kurva ROC/AUC (*Receiver Operating Characteristic/Area Under the Curve*).

1.4 TUJUAN DAN MANFAAT PENELITIAN

1.4.1 Tujuan Penelitian

Berdasarkan tinjauan yang telah dilakukan penulis, tujuan dari penelitian ini sebagai berikut :

1. Mengidentifikasi faktor-faktor risiko obesitas pada individu yang memiliki pengaruh signifikan terhadap tingkat obesitas berdasarkan data yang tersedia.
2. Mengimplementasikan dan membandingkan kinerja berbagai algoritma *Machine Learning* dalam mengklasifikasikan faktor risiko obesitas pada individu.
3. Menilai efektivitas berbagai algoritma *Machine Learning* dalam memprediksi risiko obesitas dan menentukan algoritma yang memberikan akurasi terbaik dalam klasifikasi obesitas berdasarkan faktor-faktor risiko obesitas pada individu.

1.4.2 Manfaat Penelitian

Berdasarkan tinjauan yang telah dilakukan penulis, manfaat dari penelitian ini sebagai berikut:

1. Memberikan wawasan mengenai faktor-faktor risiko yang paling berpengaruh dalam meningkatkan risiko obesitas, sehingga dapat dijadikan dasar untuk perencanaan strategi pencegahan.
2. Menyediakan referensi mengenai algoritma *Machine Learning* yang paling efektif dalam mendeteksi tingkat obesitas, yang dapat digunakan untuk pengembangan sistem pendeteksi dini.
3. Mendukung pengembangan kebijakan kesehatan yang berbasis data, dengan menyediakan sistem klasifikasi obesitas yang akurat dan dapat membantu dalam penanganan dan pencegahan obesitas secara lebih tepat sasaran.

1.5 SISTEMATIKA PENULISAN

Sistematika ini memberikan gambaran umum mengenai topik yang dibahas dalam setiap bab penelitian yang terdiri dari 5 (lima) bab. Penjelasan untuk setiap bab adalah sebagai berikut:

BAB I : PENDAHULUAN

Bab ini terdiri dari latar belakang penelitian, perumusan, masalah, batasan masalah, tujuan dan manfaat penelitian, serta sistematika penulisan dari “Klasifikasi Multikelas Faktor Risiko Obesitas Pada Individu Menggunakan Algoritma *Machine Learning*”.

BAB II : LANDASAN TEORI

Bab ini menguraikan konsep obesitas, klasifikasi, dan faktor-faktor yang mempengaruhi. Selain itu, bab ini juga menjelaskan teori-teori terkait algoritma *Machine Learning* yang akan digunakan, serta melakukan tinjauan pustaka dari penelitian-penelitian sebelumnya yang relevan dengan topik.

BAB III : METODOLOGI PENELITIAN

Bab ini menguraikan kerangka kerja penelitian atau alur penelitian, bahan dan alat penelitian, serta metode yang digunakan dalam penulisan penelitian.

BAB IV : PEMBAHASAN DAN HASIL PENELITIAN

Bab ini menyajikan hasil yang diperoleh dari analisis data serta pembahasan mendalam mengenai hasil tersebut. Hasil yang disajikan mencakup statistik deskriptif dan evaluasi dari model *Machine Learning* yang digunakan.

BAB V : PENUTUP

Bab ini mencakup kesimpulan dari hasil penelitian yang telah dilaksanakan serta memberikan saran berdasarkan temuan penelitian ini. Bab ini juga mengusulkan arah penelitian lanjutan yang dapat dilakukan di masa mendatang.

