

BAB I

PENDAHULUAN

1.1 Latar Belakang

Obesitas merupakan salah satu permasalahan kesehatan global yang terus mengalami peningkatan dari tahun ke tahun dan menjadi faktor risiko utama berbagai penyakit tidak menular seperti diabetes melitus tipe 2, penyakit kardiovaskular, hipertensi, serta gangguan metabolik lainnya (World Health Organization, 2025). Organisasi Kesehatan Dunia (World Health Organization/WHO) melaporkan bahwa lebih dari 2,5 miliar orang dewasa di dunia berada dalam kondisi kelebihan berat badan, dan sekitar 890 juta di antaranya diklasifikasikan sebagai obesitas. Kondisi ini tidak hanya berdampak pada kualitas hidup individu, tetapi juga menimbulkan beban ekonomi yang signifikan bagi sistem kesehatan global (Tremmel et al., n.d.).

Obesitas didefinisikan sebagai kondisi penumpukan lemak tubuh yang berlebihan akibat ketidakseimbangan antara asupan energi dan pengeluaran energi dalam jangka waktu tertentu (Hales et al., 2020). Penentuan tingkat obesitas umumnya dilakukan menggunakan *Body Mass Index* (BMI), yaitu rasio antara berat badan dan kuadrat tinggi badan. Berdasarkan nilai BMI, obesitas diklasifikasikan ke dalam beberapa level, mulai dari berat badan kurang (*underweight*), berat badan normal, *overweight*, hingga obesitas tingkat I, II, dan III (Consultation, 2004).

Prevalensi obesitas di Indonesia menunjukkan tren peningkatan yang signifikan dalam beberapa tahun terakhir. Kementerian Kesehatan Republik Indonesia menyatakan bahwa pada tahun 2025, sekitar satu dari empat orang dewasa di Indonesia mengalami obesitas. Berdasarkan *cut-off Body Mass Index* (BMI) yang digunakan di Indonesia ($BMI \geq 27$), prevalensi obesitas meningkat dari 10,3% pada tahun 2007 menjadi 22,3% pada tahun 2023 (Indonesia, 2019). Sementara itu, jika menggunakan standar *World Health Organization* (WHO) dengan *cut-off* $BMI \geq 30$, prevalensi obesitas juga mengalami peningkatan dari 3,7% menjadi 10,4% pada periode yang sama (Kementerian Kesehatan Republik Indonesia, 2023). Data tersebut juga

menunjukkan bahwa prevalensi obesitas lebih tinggi pada perempuan, yaitu sebesar 29,3%, dibandingkan dengan laki-laki (Nursalam et al., 2025). Kondisi ini mengindikasikan bahwa obesitas telah menjadi permasalahan kesehatan masyarakat yang serius dan memerlukan pendekatan penanganan yang lebih efektif dan berbasis data.

Perkembangan teknologi informasi dan ketersediaan data kesehatan dalam jumlah besar mendorong penerapan *machine learning* dalam bidang kesehatan, khususnya untuk klasifikasi dan prediksi tingkat obesitas. Metode *machine learning* mampu mengolah data multidimensi yang melibatkan faktor gaya hidup, kondisi fisik, serta kebiasaan makan secara lebih efektif dibandingkan metode statistik konvensional (Obermeyer & Emanuel, n.d.). Beberapa algoritma *machine learning* telah digunakan dalam penelitian klasifikasi obesitas, antara lain *Naïve Bayes*, *K-Nearest Neighbor* (KNN), *Decision Tree*, *Artificial Neural Network* (ANN), dan *Support Vector Machine* (SVM) (De-La-Hoz-Correa et al., 2019). Namun, karakteristik data obesitas yang bersifat non-linear, saling berkorelasi, serta memiliki lebih dari dua kelas (*multiclass*) menyebabkan tidak semua algoritma mampu memberikan performa optimal.

Support Vector Machine (SVM) dikenal sebagai algoritma klasifikasi yang kuat dalam menangani data berdimensi tinggi dan hubungan non-linear melalui penggunaan fungsi kernel (Cortes & Vapnik, 1995). Pada permasalahan klasifikasi obesitas, SVM sering dikombinasikan dengan kernel *Radial Basis Function* (RBF) untuk meningkatkan kemampuan pemisahan antar kelas (Razak et al., 2024). Akan tetapi, secara dasar SVM merupakan algoritma klasifikasi biner sehingga diperlukan pendekatan *multiclass* SVM, seperti *One-Against-One* (OAO) atau *One-Against-All* (OAA), untuk menangani klasifikasi dengan lebih dari dua kelas obesitas.

Dataset obesitas yang banyak digunakan dalam penelitian adalah *Obesity Dataset* yang berisi atribut gaya hidup dan kondisi fisik individu, seperti usia, tinggi badan, berat badan, kebiasaan makan, aktivitas fisik, konsumsi air, serta riwayat keluarga obesitas (Palechor & de la Hoz Manotas, 2019). Dataset ini mengklasifikasikan obesitas ke dalam tujuh kelas, yaitu *Insufficient*

Weight, Normal Weight, Overweight Level I, Overweight Level II, Obesity Type I, Obesity Type II, dan Obesity Type III. Penelitian sebelumnya menunjukkan bahwa SVM memiliki performa yang cukup baik dalam klasifikasi obesitas dibandingkan algoritma lain. Studi yang dilakukan oleh Razak et al. menunjukkan bahwa SVM dengan kernel RBF mampu mencapai tingkat akurasi sebesar 89% dalam klasifikasi obesitas berbasis gaya hidup dan kondisi fisik (Kurniabudi et al., 2022). Namun demikian, sebagian penelitian masih berfokus pada perbandingan algoritma tanpa mendalami implementasi SVM dalam skema *multiclass* secara spesifik.

Meningkatnya prevalensi obesitas serta kompleksitas faktor yang memengaruhinya menuntut adanya metode klasifikasi yang mampu memberikan hasil yang akurat, objektif, dan konsisten. Pendekatan konvensional dalam pengelompokan tingkat obesitas dinilai kurang optimal dalam menangani data dengan banyak atribut serta hubungan yang bersifat non-linear. Oleh karena itu, pemanfaatan metode *machine learning* menjadi penting untuk mendukung proses klasifikasi tingkat obesitas secara lebih efektif dan berbasis data. Berbeda dengan sebagian penelitian sebelumnya yang hanya berfokus pada perbandingan algoritma klasifikasi atau penerapan model secara umum, penelitian ini secara khusus mengkaji penerapan *Multiclass Support Vector Machine* dengan pendekatan *one-vs-one* dalam mengklasifikasikan tingkat obesitas ke dalam tujuh kelas berdasarkan data gaya hidup dan kondisi fisik individu. Kebaruan penelitian ini terletak pada fokus analisis implementasi SVM dalam skema *multiclass* yang dikombinasikan dengan tahapan *preprocessing* dan evaluasi yang terstruktur menggunakan perangkat lunak Orange Data Mining.

Hasil dari penelitian ini diharapkan dapat memberikan kontribusi ilmiah dalam pengembangan kajian penerapan *machine learning*, khususnya algoritma *Multiclass Support Vector Machine* pada permasalahan klasifikasi obesitas. Selain itu, penelitian ini diharapkan dapat memberikan kontribusi praktis sebagai dasar pengembangan sistem pendukung keputusan di bidang kesehatan untuk mendukung deteksi dini dan pengelompokan tingkat obesitas secara lebih akurat.

1.2 Rumusan Masalah

1.2.1 Identifikasi Permasalahan

Berdasarkan latar belakang yang telah diuraikan, permasalahan utama dalam penelitian ini adalah meningkatnya kasus obesitas yang memerlukan pengelompokan level obesitas secara tepat berdasarkan data kebiasaan hidup dan kondisi fisik individu. Proses klasifikasi level obesitas secara manual memiliki keterbatasan, antara lain membutuhkan waktu yang lama dan berpotensi menimbulkan kesalahan subjektif. Selain itu, dataset obesitas memiliki karakteristik yang kompleks, terdiri dari banyak atribut serta lebih dari dua kelas target, sehingga diperlukan metode klasifikasi yang mampu menangani permasalahan *multiclass* dan data yang bersifat non-linear. Oleh karena itu, diperlukan suatu pendekatan berbasis *machine learning* yang mampu mengklasifikasikan level obesitas secara otomatis, objektif, dan konsisten.

1.2.2 Metode Usulan

Pada penelitian ini, metode *Multiclass SVM* diimplementasikan menggunakan perangkat lunak Orange Data Mining. Dataset obesitas akan melalui tahapan *preprocessing* yang meliputi konversi atribut kategorikal menjadi numerik menggunakan metode *one-hot encoding* serta normalisasi data menggunakan metode *Min-Max Normalization*. Selanjutnya, proses klasifikasi dilakukan menggunakan algoritma SVM dengan kernel *Radial Basis Function* (RBF) dan strategi *multiclass one-vs-one*. Evaluasi performa model dilakukan menggunakan metode 5-Fold Stratified Cross Validation untuk mengukur tingkat akurasi dan kemampuan klasifikasi model. Metode usulan ini diharapkan mampu memberikan hasil klasifikasi level obesitas yang akurat serta dapat menjadi dasar pengembangan sistem pendukung keputusan di bidang kesehatan.

1.2.3 Pertanyaan Penelitian

Berdasarkan permasalahan dan metode usulan yang diidentifikasi, maka pertanyaan penelitian yang diajukan dalam penelitian ini adalah:

1. Bagaimana performa algoritma *Multiclass Support Vector Machine* (SVM) dengan kernel RBF dalam mengklasifikasikan level obesitas berdasarkan data kebiasaan hidup dan kondisi fisik individu?

2. Apakah penerapan *preprocessing* data berupa *one-hot encoding* dan *Min-Max Normalization* dapat meningkatkan akurasi klasifikasi level obesitas menggunakan *Multiclass SVM*?
3. Bagaimana hasil evaluasi kinerja model *Multiclass SVM* menggunakan metode *5-Fold Stratified Cross Validation* pada dataset obesitas?

1.3 Tujuan dan Manfaat Penelitian

1.3.1 Tujuan Penelitian

Berdasarkan permasalahan, solusi, dan pertanyaan penelitian, maka tujuan penelitian ini adalah sebagai berikut:

1. Menerapkan algoritma *Multiclass Support Vector Machine* (SVM) dengan kernel *Radial Basis Function* (RBF) untuk mengklasifikasikan level obesitas berdasarkan data kebiasaan hidup dan kondisi fisik individu menggunakan perangkat lunak Orange Data Mining.
2. Menganalisis kinerja model *Multiclass SVM* dalam menangani permasalahan *multiclass* dan karakteristik data yang bersifat non-linear pada dataset obesitas.
3. Mengukur tingkat akurasi dan kemampuan klasifikasi model melalui evaluasi menggunakan metode *5-Fold Stratified Cross Validation* dan *confusion matrix*.

1.3.2 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi sebagai berikut:

a. **Manfaat Teoritis**

Memberikan kontribusi keilmuan di bidang *machine learning*, khususnya dalam penerapan algoritma *Multiclass Support Vector Machine* (SVM) untuk klasifikasi level obesitas, serta memperkaya kajian tentang efektivitas kernel RBF dalam menangani data kesehatan yang bersifat non-linear.

b. **Manfaat Praktis**

Menyediakan alternatif proses klasifikasi level obesitas yang lebih objektif, akurat, dan konsisten sehingga dapat menjadi dasar pengembangan sistem pendukung keputusan di bidang kesehatan, baik

untuk tenaga medis maupun masyarakat umum dalam melakukan deteksi dini risiko obesitas.

c. Manfaat Metodologis

Mendemonstrasikan implementasi *Multiclass SVM* dengan menggunakan Orange Data Mining sebagai perangkat lunak *visual programming* yang mudah direplikasi, sehingga dapat diadopsi untuk keperluan edukasi maupun penelitian lanjutan.

1.4 Batasan Masalah

Agar penelitian ini lebih terarah dan tidak menyimpang dari tujuan yang diharapkan, maka diberikan batasan masalah sebagai berikut:

1. Dataset yang digunakan adalah *Obesity Dataset (Estimation of Obesity Levels Based on Eating Habits and Physical Condition)* yang bersumber dari *UCI Machine Learning Repository* dan *Kaggle*, dengan jumlah 2.111 *record* dan 17 atribut.
2. Klasifikasi obesitas dibagi menjadi tujuh kelas sesuai dengan atribut kelas *NObesyedad*, yaitu *Insufficient Weight*, *Normal Weight*, *Overweight Level I*, *Overweight Level II*, *Obesity Type I*, *Obesity Type II*, dan *Obesity Type III*.
3. Metode klasifikasi yang digunakan adalah *Support Vector Machine (SVM)* dengan pendekatan *multiclass one-vs-one*.
4. Kernel yang digunakan pada SVM adalah *Radial Basis Function (RBF)* dengan parameter *C* dan *gamma* menggunakan nilai default pada Orange Data Mining.
5. Evaluasi performa model dilakukan menggunakan *confusion matrix* dan nilai akurasi dengan metode *5-Fold Stratified Cross Validation*.
6. Perangkat lunak yang digunakan adalah Orange Data Mining tanpa pemrograman manual (*visual programming*).
7. Penelitian tidak melakukan *hyperparameter tuning* secara ekstensif dan tidak menggunakan teknik *oversampling* seperti *SMOTE*.

1.5 Sistematika Penulisan

Pada penelitian ini sistematika penulisan skripsi yang digunakan adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi latar belakang masalah, identifikasi permasalahan, metode usulan, pertanyaan penelitian, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan skripsi.

BAB II KAJIAN PUSTAKA

Bab ini membahas landasan teori yang mendukung penelitian, meliputi konsep obesitas, faktor-faktor penyebab obesitas, klasifikasi level obesitas, *machine learning* dalam bidang kesehatan, algoritma *Support Vector Machine* (SVM), *Multiclass SVM*, Orange Data Mining, *confusion matrix*, serta penelitian-penelitian terdahulu yang relevan dengan topik penelitian.

BAB III METODE PENELITIAN

Bab ini menguraikan secara rinci metodologi penelitian yang digunakan, meliputi jenis dan pendekatan penelitian, ruang lingkup dan objek penelitian, bahan dan alat penelitian, tempat dan waktu penelitian, teknik pengumpulan data, definisi operasional variabel penelitian, rancangan kegiatan penelitian, serta teknik analisis data yang diimplementasikan menggunakan Orange Data Mining.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil penelitian secara komprehensif meliputi implementasi skema penelitian pada Orange Data Mining, hasil eksplorasi dan *preprocessing* data, hasil pengujian klasifikasi dengan *Multiclass SVM* pada skenario 5-Fold Cross Validation dan *Split Data* 80:20, hasil pengujian perbandingan kernel, analisis fitur penting, analisis kesalahan klasifikasi, perbandingan dengan penelitian terdahulu, serta pembahasan interpretasi hasil dan keterbatasan penelitian.

BAB V PENUTUP

Bab ini berisi kesimpulan dari hasil penelitian yang menjawab pertanyaan penelitian serta saran untuk pengembangan penelitian selanjutnya.

DAFTAR PUSTAKA

Berisi daftar referensi yang digunakan dalam penelitian, terdiri dari jurnal ilmiah, prosiding, buku, laporan resmi, dan sumber terpercaya lainnya.

LAMPIRAN

Berisi data pendukung seperti cuplikan dataset, skema Orange Data Mining, hasil output pengujian lengkap, dokumentasi kegiatan penelitian, dan riwayat hidup penulis.