

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kanker payudara merupakan salah satu jenis kanker yang paling umum menyerang wanita di seluruh dunia dan merupakan penyebab utama kematian akibat kanker pada wanita. Menurut World Health Organization (WHO) tahun 2020, sekitar 2,3 juta wanita di seluruh dunia telah terdiagnosis kanker payudara [1]. Angka ini menunjukkan bahwa deteksi dini dan diagnosis yang tepat sangat penting untuk meningkatkan peluang kesembuhan dan menurunkan tingkat kematian akibat kanker ini.

Dengan kemajuan teknologi, metode komputasi seperti pembelajaran mesin kini banyak digunakan dalam bidang medis, termasuk dalam membantu proses klasifikasi dan diagnosis kanker [2]. Pembelajaran mesin memungkinkan komputer untuk belajar dari data dan membuat prediksi berdasarkan pola yang ditemukan, sehingga dapat digunakan sebagai alat bantu dalam pengambilan keputusan medis.

Algoritma klasifikasi sangat penting untuk klasifikasi kanker payudara karena dapat membantu menganalisis data medis dan membedakan jenis kanker, seperti membedakan kanker jinak dan ganas [3]. Dua algoritma klasifikasi yang paling umum digunakan adalah Decision Tree dan Support Vector Machine (SVM).

Untuk memisahkan kelas data dengan karakteristik berdimensi tinggi dan batas keputusan yang kompleks, algoritma Support Vector Machine (SVM) dikenal memiliki kinerja yang sangat baik [4]. SVM bekerja dengan mencari jarak terbesar antara dua kelas dan membentuk hyperplane atau batas keputusan yang ideal. Decision Tree, di sisi lain, menawarkan metode yang berbeda dengan membuat struktur pohon berdasarkan aturan yang dibuat dari data. Decision Tree memiliki banyak keuntungan, salah satunya adalah bahwa hasilnya dapat diinterpretasikan dengan mudah dan proses pengambilan keputusan dapat dilihat secara visual [5].

Namun, karakteristik dataset, seperti jumlah fitur, ketidakseimbangan kelas, dan tingkat *noise*, dapat memengaruhi seberapa efektif masing-masing algoritma [6]. Oleh karena itu, untuk klasifikasi kanker payudara, sangat penting untuk

melakukan analisis perbandingan kinerja antara algoritma Decision Tree dan SVM. Analisis ini akan membandingkan kinerja kedua algoritma berdasarkan metrik evaluasi seperti akurasi, presisi, recall, dan skor F1. Tujuan dari penelitian ini adalah untuk menentukan algoritma mana yang lebih efisien dan efektif dalam mengklasifikasikan data kanker payudara.

Hasil penelitian ini diharapkan dapat memberikan kontribusi dalam bidang machine learning, khususnya dalam menganalisis dan membandingkan performa algoritma klasifikasi untuk kasus diagnosis kanker payudara, serta memberikan rekomendasi algoritma yang paling optimal berdasarkan metrik evaluasi yang komprehensif.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana hasil analisis kinerja algoritma Support Vector Machine (SVM) dalam klasifikasi kanker payudara berdasarkan dataset yang digunakan?
2. Bagaimana hasil analisis kinerja algoritma Decision Tree dalam klasifikasi kanker payudara berdasarkan dataset yang sama?
3. Algoritma manakah yang menunjukkan kinerja yang lebih optimal dalam klasifikasi kanker payudara, jika dibandingkan berdasarkan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score?

1.3 Batasan Masalah

Agar penelitian ini terfokus dan terarah, maka batasan masalah dalam penelitian ini ditentukan sebagai berikut:

1. Penelitian ini hanya membandingkan dua algoritma klasifikasi, yaitu Support Vector Machine (SVM) dan Decision Tree, tanpa melibatkan algoritma lainnya.

2. Dataset yang digunakan dalam penelitian ini adalah dataset kanker payudara yang bersumber dari UCI Machine Learning Repository, yaitu Wisconsin Breast Cancer Dataset (WBCD).
3. Penelitian ini berfokus pada proses klasifikasi kanker payudara ke dalam dua kategori, yaitu jinak (benign) dan ganas (malignant).
4. Evaluasi kinerja algoritma hanya dilakukan berdasarkan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score.
5. Penelitian ini tidak membahas aspek medis atau biologis dari kanker payudara secara mendalam, melainkan terbatas pada implementasi dan evaluasi algoritma machine learning.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Menganalisis kinerja algoritma Support Vector Machine (SVM) dalam melakukan klasifikasi kanker payudara.
2. Menganalisis kinerja algoritma Decision Tree dalam melakukan klasifikasi kanker payudara.
3. Membandingkan hasil klasifikasi antara algoritma SVM dan Decision Tree berdasarkan metrik evaluasi seperti akurasi, precision, recall, F1-score, dan AUC-ROC untuk menentukan algoritma yang paling optimal.

1.5 Manfaat Penelitian

Adapun manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Manfaat Akademis
Memberikan kontribusi dalam pengembangan ilmu pengetahuan di bidang machine learning, khususnya dalam penerapan algoritma klasifikasi untuk data medis.
2. Manfaat Praktis
Memberikan informasi dan rekomendasi mengenai algoritma yang paling efektif dan akurat untuk digunakan dalam sistem pendukung keputusan diagnosis kanker payudara.

3. Manfaat Teknologis

Menjadi dasar pengembangan aplikasi berbasis kecerdasan buatan yang dapat membantu tenaga medis dalam mendeteksi kanker payudara secara cepat dan akurat.

4. Manfaat Sosial

Mendukung upaya peningkatan kualitas layanan kesehatan dengan menyediakan pendekatan analitik berbasis data untuk deteksi dini kanker payudara.

