

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Perkembangan teknologi informasi dan kecerdasan buatan memberikan peluang besar dalam meningkatkan efisiensi dan akurasi dalam proses diagnosis penyakit. Salah satu cabang dari kecerdasan buatan yang banyak digunakan dalam bidang medis adalah *machine learning*, terutama teknik klasifikasi. Dengan memanfaatkan data klinis dan demografis pasien, algoritma klasifikasi dapat digunakan untuk memprediksi kemungkinan adanya kanker paru-paru tanpa harus melakukan prosedur invasif atau mahal. Hal ini tentu saja sangat membantu dalam proses skrining awal yang cepat dan bersifat non-invasif (Zaenudin & Riyan, 2024).

Teknologi informasi dan komunikasi telah berkembang dengan cepat, menghasilkan jumlah data yang sangat besar. Data-data tersebut perlu dianalisis dan diatur agar dapat dimanfaatkan secara optimal. Dalam hal ini, klasifikasi memegang peranan penting dalam mengelompokkan data serta mengidentifikasi pola yang bermanfaat. Klasifikasi adalah proses membangun sekumpulan model atau fungsi yang menggambarkan perbedaan antar kategori data, dengan tujuan untuk memprediksi kategori objek yang belum diketahui. Model ini diperoleh melalui analisis data yang memiliki label kelas yang sudah diketahui (Ramadhani dkk., 2022).

Beberapa algoritma klasifikasi populer yang banyak digunakan dalam penelitian medis adalah *Support Vector Machine* (SVM) dan *Decision Tree* (DT). Algoritma SVM memiliki keunggulan dalam mengolah data berdimensi tinggi serta mampu membentuk batas klasifikasi yang optimal antara dua kelas data. Sementara itu, *Decision Tree* dikenal mudah diinterpretasikan dan cocok digunakan dalam kasus-kasus di mana pemahaman logika klasifikasi menjadi penting (Amin, 2024). Kedua algoritma ini memiliki karakteristik dan kelebihan masing-masing, sehingga penting untuk dilakukan studi komparatif guna mengetahui mana yang paling cocok untuk digunakan dalam konteks deteksi kanker paru-paru. Deteksi dini kanker paru-paru sangat penting untuk

meningkatkan peluang kesembuhan pasien. Namun, dalam praktiknya, proses diagnosis sering kali memerlukan waktu yang lama dan bergantung pada keahlian medis yang terbatas. Oleh karena itu, pemanfaatan metode klasifikasi dalam *machine learning* menjadi salah satu pendekatan potensial untuk membantu mendiagnosis penyakit ini secara lebih cepat dan akurat.

Berbagai algoritma klasifikasi telah dikembangkan, termasuk *Support Vector Machine* (SVM) dan *Decision Tree* (DT). Meskipun kedua metode ini banyak digunakan dalam berbagai studi klasifikasi medis, hasil klasifikasinya dapat bervariasi tergantung pada karakteristik data yang digunakan. SVM dikenal memiliki performa tinggi pada data berdimensi besar, sementara *Decision Tree* unggul dalam interpretasi hasil.

Dalam penelitian ini akan membandingkan hasil klasifikasi SVM dan DT dalam deteksi pasien kanker paru-paru. Penyakit kanker paru-paru sering kali tidak menimbulkan gejala pada tahap awal sehingga deteksi dini atau pemekrisaan rutin sangat penting dilakukan, terutama bagi kelompok berisiko tinggi seperti perokok atau mantan perokok. Jika kanker ditemukan lebih dini sebelum menyebar ke organ lain, maka pemilihan pengobatan lebih banyak dan lebih efektif sehingga peluang kesembuhan lebih besar. Deteksi dini memungkinkan penanganan lebih cepat sehingga pasien memiliki kemungkinan hidup lebih lama dan kualitas hidup yang lebih baik. Perbandingan hasil klasifikasi antara metode *Support Vector Machine* dan *decision Tree* dalam deteksi kanker paru-paru dilakukan untuk mengetahui metode klasifikasi mana yang lebih optimal dalam menghasilkan prediksi yang akurat dan dapat diandalkan. Masalah yang muncul adalah belum diketahui secara pasti metode mana yang lebih optimal dalam mengklasifikasikan data kanker paru-paru dengan akurasi dan efisiensi terbaik. Oleh karena itu, perlu dilakukan penelitian untuk membandingkan kedua metode tersebut berdasarkan metrik evaluasi seperti akurasi, presisi, dan *recall*, guna menentukan algoritma yang paling tepat digunakan dalam konteks ini.

Pada penelitian sebelumnya, penulis Dhini Septhya, Kharisma Rahayu, Salsabila Rabbani, Vindi Fitria, Rahmaddeni, Yuda Irawan, dan Regiolina Hayami dengan judul penelitian Implementasi Algoritma *Decision Tree* dan *Support Vector Machine* Untuk Klasifikasi Penyakit Kanker Paru, penelitian ini

menunjukkan algoritma SVM memiliki akurasi yang lebih unggul yaitu 62%. Pada penelitian terdahulu yang ditulis oleh Nia Nuraeni dan Puji Astuti dengan judul penelitian pendekatan Machine Learning untuk Deteksi Dini Kanker Paru-Paru: Mengoptimalkan Sensitivitas dan Akurasi menunjukkan metode SVM memiliki akurasi tinggi 89% dan *Decision Tree* mencatat akurasi tertinggi 90%. Pada Penelitian lain yang ditulis oleh Sri Indra Maiyanti<sup>1</sup>, Des Alwine Zayanti, Yuli Andriani, Bambang Suprihatin, Anita Desiani, Aulia Salsabila, Nyayu Chika Marselina dengan judul penelitian Perbandingan Klasifikasi Penyakit Kanker Paru-paru menggunakan *Support Vector Machine* dan *K-Nearest Neighbor* menunjukkan hasil akurasi tertinggi diperoleh oleh metode SVM dengan nilai 95,16%.

Berdasarkan latar belakang permasalahan diatas dapat diangkat sebuah judul “Perbandingan Hasil Klasifikasi Dalam Deteksi Pasien Kanker Paru-Paru Menggunakan Metode *Support Vector Machine* dan *Decision Tree*”. Penelitian ini bertujuan untuk mencari performansi hasil klasifikasi yang terbaik pada *dataset* kanker paru-paru yang diambil dari sumber *kaggle.com* dan RSUD dr.Zubir Mahmud Kecamatan Aceh Timur. Penelitian ini diharapkan dapat memberikan wawasan yang berharga peneliti dan praktisi dalam pemodelan klasifikasi serta membantu dalam memilih metode yang paling sesuai untuk deteksi pasien kanker paru-paru.

## **1.2 Rumusan Masalah**

Berdasarkan penjelasan sebelumnya, penulis berupaya merumuskan masalah dengan cermat agar penelitian yang dilakukan lebih terarah dan dapat mencapai tujuan yang diharapkan. Dalam penelitian ini, rumusan masalah yang diajukan adalah sebagai berikut:

1. Bagaimana membuat aplikasi perbandingan hasil klasifikasi dalam deteksi kanker paru-paru menggunakan metode *Support Vector Machine* dan *Decision Tree* ?
2. Bagaimana membandingkan hasil klasifikasi dalam deteksi pasien kanker paru-paru dengan metode *Support Vector Machine* dan *Decision Tree* ?

### **1.3 Tujuan Penelitian**

Adapun tujuan yang ingin dicapai dari penelitian ini adalah sebagai berikut:

1. Membuat aplikasi perbandingan hasil klasifikasi dalam deteksi kanker paru-paru menggunakan metode *Support Vector Machine* dan *Decision Tree*.
2. Membandingkan hasil klasifikasi dalam deteksi pasien kanker paru-paru dengan metode *Support Vector Machine* dan *Decision Tree*.

### **1.4 Manfaat Penelitian**

Berikut adalah manfaat yang dapat diambil dari penelitian ini:

1. Dengan membandingkan hasil klasifikasi kedua metode, penelitian ini dapat mengidentifikasi algoritma yang memberikan hasil klasifikasi yang lebih tinggi dalam deteksi penyakit kanker paru-paru.
2. Pengembangan model klasifikasi yang lebih baik, dengan memahami kelebihan dan kekurangan SVM dan *Decision Tree* dalam konteks deteksi kanker paru-paru, penelitian ini dapat membantu dalam pengembangan model klasifikasi yang lebih baik dan lebih akurat.
3. Hasil dari perbandingan ini dapat membantu dalam pengembangan teknologi deteksi dini penyakit kanker paru-paru berbasis *machine learning*, yang dapat memberikan manfaat besar dalam upaya pencegahan dan pengobatan penyakit tersebut.

### **1.5 Ruang Lingkup dan Batasan Penelitian**

Dengan mempertimbangkan latar belakang dan rumusan masalah yang sudah dijelaskan, penulis membuat batasan masalah sebagai berikut:

1. Penelitian ini terbatas pada penggunaan dataset yang diambil dari *Kaggle* dan dari RSUD dr.Zubir Mahmud Kabupaten Aceh Timur terkait dengan pasien penyakit kanker paru-paru, dengan jumlah data sebanyak 361 *record*.
2. *Output* yang dihasilkan dari masing-masing algoritma klasifikasi adalah yang terkena kanker/tidak.
3. Fitur yang digunakan adalah faktor resiko dan gejala dari penyakit yang derita.

4. Parameter data terdiri dari *gender, age, smoking, yellow finger, anxiety, peer pressure, chonic disease, fatigue, allergy, wheezing, alcohol consuming, coughing, shortness of breath, swallowing difficulty, chest pain* dan *lung cancer*.
5. Fokus penelitian ini adalah pada perbandingan hasil klasifikasi *Support Vector Machine* dan *Decision Tree* untuk klasifikasi pasien penyakit kanker paru-paru.
6. Program ini menggunakan nilai evaluasi dari *Accuracy, Precision* dan *Recall* menggunakan *convusion matrix*.