

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Mata sebagai salah satu indra utama manusia memegang peranan penting dalam menjalankan aktivitas sehari-hari. Fungsi utama mata adalah sebagai mekanisme pendeteksi cahaya, memungkinkan manusia membedakan antara terang dan gelap serta membentuk persepsi visual, yang memungkinkan kita melihat sekeliling. Retina merupakan salah satu komponen terpenting pada mata karena mengubah impuls cahaya menjadi sinyal saraf, yang kemudian diinterpretasikan oleh otak sebagai gambar yang kita lihat. Proses ini mendasari kemampuan visual manusia. Namun bila ada penyakit mata, dapat mengganggu fungsi penglihatan dan menyulitkan melakukan aktivitas normal. Masalah kesehatan mata, seperti katarak, glaukoma, dan retinopati diabetik, dapat berkisar dari ringan hingga berat. Katarak, misalnya, terjadi ketika lensa mata menjadi keruh sehingga mengganggu masuknya cahaya dan mengakibatkan penglihatan kabur. Hal ini dapat disebabkan oleh berbagai faktor seperti oksidasi, paparan radiasi ultraviolet, keturunan, dan nutrisi.

Diagnosa yaitu proses pengidentifikasian kelemahan atau kondisi medis yang dialami seseorang melalui pemeriksaan dan analisis cermat terhadap gejala-gejala yang muncul (Iskandar, 2020). Dalam proses mengidentifikasi penyakit, biasanya dilakukan diagnosa tidak langsung.

*Data mining* merupakan proses pengolahan data untuk memperoleh informasi tambahan yang sebelumnya tidak terdeteksi secara manual, dan akan disimpan ke sebuah database khusus. Proses ini memerlukan ekstraksi dari pola-pola penting dari data yang tersedia untuk membentuk model atau fungsi yang menggambarkan serta membedakan konsep atau label data (Ismanto & Novalia, 2021). Teknik *data mining* banyak digunakan diberbagai industri. Penerapan data mining di sektor kesehatan telah menjadi model yang baru dan populer, karena dengan menggunakan teknik ini, kita dapat mengekstraksi informasi dan menemukan pola-pola tersembunyi dalam data yang dapat digunakan sebagai

dasar untuk pengambilan keputusan. Salah satu metode dalam data mining yang umum digunakan adalah teknik klasifikasi, yang bertujuan untuk memprediksi variabel-variabel yang terkait. Algoritma klasifikasi seperti *naive bayes*, *decision tree*, *random forest*, dan *support vector machine* adalah contoh-contoh dari algoritma yang digunakan dalam proses ini. Dengan menggunakan algoritma klasifikasi ini, kita dapat melakukan prediksi awal terhadap kondisi seorang pasien, yang dapat membantu dalam upaya pencegahan kematian dini.

Klasifikasi, suatu bentuk analisis data, membantu dalam memprediksi label kelas untuk sampel yang diberikan. Berbagai teknik klasifikasi telah muncul di berbagai bidang seperti pembelajaran mesin, sistem pakar, dan statistik. Sejak tahun 1990-an, para peneliti telah mengkurasi repositori perangkat lunak untuk meningkatkan pemahaman mereka tentang data. Untuk memulai sebuah klasifikasi dibutuhkan suatu metode data mining. Pada penelitian ini menggunakan algoritma ID3 yang digunakan untuk membuat pohon keputusan yang menggunakan konsep dari entropi information. Algoritma ID3 tersebut dapat menjadi alternatif pilihan yang tepat, namun sampai saat ini belum diketahui algoritma yang paling akurat dalam diagnosa penyakit mata katarak tersebut.

Teknik *random forest* dibangun di atas pendekatan *decision tree* dengan menggunakan beberapa *decision tree*, dengan setiap *decision tree* dilatih pada sampel yang berbeda dan memisahkan atribut yang dipilih secara acak dari subset. Teknik ini menawarkan beberapa manfaat, seperti meningkatkan akurasi dengan adanya data yang hilang dan menangani *outlier* secara efektif, serta efisien dalam penyimpanan data. Selain itu, metode ini juga menggabungkan mekanisme pemilihan fitur untuk mengidentifikasi fitur yang optimal, sehingga meningkatkan kinerja model klasifikasi (Supriyadi et al., 2020).

Menurut Kurnia *et al.*, (2019) “Penerapan Algoritma *K-Means* Untuk Mengelompokkan Diagnosa Penyakit Mata Berdasarkan Rentang Usia” tujuan dari penelitian ini adalah untuk menemukan informasi dari kumpulan data rekam medis yang terkait dengan penyakit mata. Temuan perbandingan menunjukkan keseragaman. Awalnya, metode perbandingan *Between-Class Variation* (BCV) dan *Within-Class Variation* (WCV) digunakan. Rasio BCV dan WCV adalah

0,002 yang menunjukkan kualitas tinggi dari pemanfaatan nilai *centroid*. Selanjutnya, akurasi hasil pengelompokan dinilai dengan menggunakan metode *Receiver Operating Characteristic* (ROC) dengan nilai 0,645 yang menunjukkan tingkat akurasi yang memuaskan. Pada tahap selanjutnya, perbandingan dilakukan antara hasil aplikasi dan hasil dari *rapidminer*. Perbandingan pengujian menunjukkan bahwa hasil *cluster* tidak melebihi 3% untuk tiap *cluster*. Secara khusus, *cluster* 1 tidak menunjukkan adanya perbedaan (0%) pada data, sedangkan *cluster* 2 dan 3 menunjukkan perbedaan masing-masing 108 data (2,2%).

Menurut Prandika Siregar *et al.*, (2023) “Implementasi Algoritma *Random Forest* Dalam Klasifikasi Diagnosa Penyakit Stroke” penggunaan Metode *Random Forest* diharapkan sebagai pilihan optimal untuk preprocessing data dalam mendeteksi gejala awal. Penyesuaian model menghasilkan skor pelatihan 96% sedangkan hasil *precision*, *recall*, *F1-score*, dan akurasi, yang digambarkan dalam table, menunjukkan tingkat akurasi sebesar 95%. Selain itu, hasil AUC sebesar 0,80 menegaskan kualitas klasifikasi model yang memuaskan.

Penelitian ini menerapkan metode *random forest* dalam diagnosa penyakit mata katarak. Metode ini telah terbukti efektif dalam mengelompokkan data medis dan memprediksi kondisi kesehatan, dengan menggabungkan beberapa *decision tree* yang dilatih pada sampel berbeda. *Random forest* mampu meningkatkan akurasi diagnosa dengan menangani data yang hilang dan outlier secara efektif, serta mengidentifikasi fitur optimal untuk meningkatkan kinerja model klasifikasi. Studi terdahulu menunjukkan bahwa penggunaan teknik ini berhasil dalam berbagai konteks kesehatan, termasuk deteksi dini gejala penyakit.

Berdasarkan penjelasan diatas, maka penelitian yang akan dilakukan adalah **“DIAGNOSA PENYAKIT MATA KATARAK MENGGUNAKAN METODE *RANDOM FOREST* (RF)”**.

## 1.2 Rumusan Masalah

Berdasarkan uraian latar belakang diatas rumusan masalah pada penelitian ini adalah sebagai berikut :

1. Bagaimana merancang sistem diagnosa penyakit mata katarak?
2. Bagaimana menerapkan metode *random forest* dalam diagnosa penyakit mata katarak?

### 1.3 Batasan Masalah

Adapun batasan masalah yang ditetapkan pada penelitian ini adalah sebagai berikut :

1. Penelitian ini difokuskan pada studi kasus yang dilakukan di Rumah Sakit Umum Daerah (RSUD) Muda Sedia Aceh Tamiang. Pemilihan lokasi ini dilakukan karena institusi tersebut memiliki data yang relevan dan mendukung penelitian terkait deteksi risiko katarak.
2. Data yang digunakan dalam penelitian ini bersumber dari rekam medis pasien RSUD Aceh Tamiang selama periode tahun 2019 hingga 2023. Data tersebut mencakup berbagai variabel yang relevan untuk mendeteksi risiko katarak.
3. Penelitian ini secara khusus difokuskan pada salah satu jenis gangguan penglihatan, yaitu penyakit mata katarak. Katarak dipilih sebagai objek penelitian karena merupakan salah satu penyebab utama kebutaan yang dapat dideteksi secara dini melalui analisis gejala dan pemeriksaan kesehatan mata.
4. Variabel-variabel yang dianalisis dalam penelitian ini meliputi gejala klinis yang sering dialami oleh pasien dengan risiko katarak. Gejala tersebut mencakup usia, riwayat keluarga, penurunan ketajaman penglihatan, kekeruhan lensa, sensitivitas terhadap cahaya, perubahan persepsi warna, munculnya penglihatan ganda, dan pandangan yang tetap buram meskipun ukuran kacamata telah disesuaikan. Variabel tambahan seperti tekanan intraokular, hasil slit lamp, dan visus juga dimasukkan dalam analisis untuk meningkatkan akurasi deteksi.
5. Penelitian ini menggunakan algoritma *Random Forest* sebagai metode klasifikasi utama. Algoritma ini dipilih karena memiliki performa yang

tinggi dalam mengolah data dengan jumlah variabel yang banyak dan memberikan hasil yang akurat dalam prediksi risiko katarak.

6. Untuk implementasi dan pengembangan sistem, penelitian ini menggunakan bahasa pemrograman Python. Bahasa ini dipilih karena mendukung berbagai pustaka pembelajaran mesin seperti *scikit-learn* dan *pandas*, yang sangat membantu dalam proses analisis data, pelatihan model, dan visualisasi hasil.

#### **1.4 Tujuan Penelitian**

Berdasarkan latar belakang dan rumusan masalah di atas maka tujuan penelitian ini adalah sebagai berikut :

1. Untuk merancang sistem diagnosa penyakit mata katarak.
2. Untuk menerapkan metode *random forest* dalam diagnosa penyakit mata katarak.

#### **1.5 Manfaat Penelitian**

Adapun manfaat dari penelitian ini adalah sebagai berikut :

1. Memberikan kemudahan bagi pasien untuk mendapatkan informasi tentang penyakit mata berdasarkan gejala yang dialaminya.
2. Dapat menjaga kesehatan mata agar mata pasien tidak mudah lelah yang menyebabkan terjadinya sakit mata.
3. Sebagai sarana untuk menerapkan penelitian yang akan diperoleh selama menempuh masa studi Strata-1.