

# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Hati adalah organ vital manusia yang memiliki fungsi kompleks dan beragam, Salah satu penyakit yang menyerang hati adalah *hepatitis* atau *liver*. Menurut data *WHO (World Health Organization)* Global Hepatitis Report 2017 menampilkan hampir 1,2 juta orang per tahun khususnya di Asia Tenggara dan Afrika mengalami kematian akibat terserang *penyakit liver* [1]. Penyakit hati merupakan salah satu penyebab utama *morbiditas dan mortalitas global*. Terdapat berbagai jenis penyakit hati, seperti perlemakan hati (*fatty liver*), *sirosis*, *hepatitis*, *kanker hati*, dan *penyakit hati kronis*. Salah satu penyebab umum dari perlemakan hati adalah akumulasi lemak *trigliserida* yang berlebihan, yang dapat berujung pada komplikasi serius seperti *sirosis* atau *kanker hati* [2].

Dalam penelitian ini, *penyakit liver* diklasifikasikan berdasarkan parameter medis tertentu yang digunakan dalam dataset penelitian. Dataset yang digunakan memuat berbagai informasi medis pasien, seperti kadar *enzim hati (ALT, AST)*, *kadar bilirubin*, *albumin*, *total protein*, serta *faktor risiko lain* yang berhubungan dengan kondisi kesehatan hati [3]. Data ini akan digunakan sebagai variabel input dalam proses klasifikasi untuk menentukan apakah seseorang mengalami *penyakit liver* atau tidak. Selain itu, *algoritma Random Forest* digunakan dalam penelitian ini untuk melakukan *klasifikasi pasien* berdasarkan parameter medis yang tersedia. Akurasi model akan diukur berdasarkan metrik evaluasi seperti *akurasi*, *precision*, *recall*, dan *F1-score*, guna mengetahui seberapa baik model dapat memprediksi penyakit liver [4].

Penelitian ini mengacu pada konsep *machine learning*, khususnya dalam bidang *data mining dan klasifikasi*. *Data mining* adalah proses ekstraksi informasi dari kumpulan data yang besar untuk menemukan pola atau hubungan yang tersembunyi. Salah satu teknik utama dalam data mining adalah *klasifikasi*, yaitu metode untuk mengelompokkan data ke dalam kategori tertentu berdasarkan pola

yang ditemukan dalam *data historis* [5]. *Algoritma Random Forest* merupakan salah satu metode klasifikasi berbasis pohon keputusan (*Decision Tree*), di mana kumpulan pohon keputusan (*ensemble learning*) digunakan untuk meningkatkan akurasi prediksi. Algoritma ini memiliki beberapa keunggulan, seperti ketahanan terhadap data yang hilang (*missing values*), kemampuan menangani data dengan banyak variabel, serta performa yang baik dalam menghadapi *outlier* [6].

Penelitian ini dilakukan untuk menerapkan algoritma *Random Forest* yang telah di optimasi dengan *Random Search* untuk memprediksi *penyakit liver* guna meningkatkan efisiensi dan akurasi dalam diagnosis penyakit tersebut [7]. Dataset yang digunakan dalam penelitian ini berasal dari *Kaggle*, yang berisi informasi medis terkait kondisi pasien dengan dugaan *penyakit liver*. Diharapkan hasil penelitian ini dapat meningkatkan akurasi deteksi dini *penyakit liver*, sehingga dapat memberikan manfaat dalam *sistem pendukung keputusan di bidang medis*. Dengan pendekatan berbasis *machine learning*, model yang dikembangkan dapat digunakan sebagai alat bantu dalam analisis medis untuk meningkatkan kecepatan dan ketepatan dalam mendeteksi *penyakit liver* [8].

### 1.2 Rumusan Masalah

Berdasarkan latar belakang masalah di atas, rumusan masalah dalam penelitian ini adalah Bagaimana kinerja *Algoritma Random Forest* yang telah di Optimasi dengan *Random Search* dalam mendeteksi dini *penyakit Liver*.

### 1.3 Batasan Masalah

Agar penelitian ini lebih terarah dan fokus, maka terdapat beberapa batasan masalah sebagai berikut:

1. Data yang digunakan dalam penelitian ini berasal dari dataset yang telah tersedia secara publik mengenai *penyakit liver*.
2. Model prediksi yang digunakan dalam penelitian ini hanya menggunakan *Algoritma Random Forest*.
3. Parameter evaluasi yang digunakan untuk mengukur performa model adalah *akurasi, presisi, recall, dan F1-score*.

4. Penelitian ini tidak membahas aspek medis secara mendalam, melainkan lebih berfokus pada penerapan dan evaluasi model machine learning.

#### 1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk menyelesaikan permasalahan yang telah diidentifikasi dan mencapai hasil yang dapat diukur ketercapaiannya. Adapun tujuan spesifik yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Menerapkan *Algoritma Random Forest* dalam prediksi *penyakit liver*.
2. Mengukur tingkat akurasi model yang diterapkan dalam *klasifikasi penyakit liver*.
3. Mengidentifikasi faktor-faktor utama yang berkontribusi dalam diagnosis *penyakit liver* berdasarkan model yang dikembangkan.

#### 1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. **Manfaat Teoritis:** Menambah wawasan dalam bidang kecerdasan buatan, khususnya dalam penerapan *Algoritma Random Forest* untuk analisis data medis.
2. **Manfaat Praktis:** Memberikan solusi bagi tenaga medis dalam membantu proses deteksi dini *penyakit liver* secara lebih cepat dan efisien.
3. **Manfaat bagi Pengembangan Sistem:** Dapat menjadi referensi dalam pengembangan sistem pendukung keputusan berbasis *machine learning* dalam bidang kesehatan.

#### 1.6 Sistematika Penulisan

Berikut adalah sistematika penulisan pada penelitian ini:

### BAB I PENDAHULUAN

Pada bab ini menjelaskan mengenai latar belakang, rumusan masalah, Batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

### BAB II TINJAUAN PUSTAKA

Bab ini berisi pembahasan dasar teori yang berkaitan dengan penelitian serta penelitian-penelitian sebelumnya yang relevan.

### **BAB III METODE PENELITIAN**

Bab ini menguraikan tahapan yang akan dilaksanakan dalam penelitian. Setiap tahap perencanaan penelitian akan dijabarkan secara terperinci berdasarkan suatu kerangka kerja yang telah ditetapkan. Selain itu, bab ini juga mencakup perancangan manajemen proyek untuk pelaksanaan penelitian, yang bertujuan untuk kelancaran dan keberhasilan seluruh proses.

### **BAB IV HASIL DAN PEMBAHASAN**

Bab ini berisi hasil dari penerapan model yang dijelaskan pada bab sebelumnya. Hasil optimasi metode *Random Search* pada *algoritma Random Forest* serta metrik evaluasi model akan dibahas secara rinci. Bab ini juga membandingkan hasil sebelum dan sesudah dilakukan optimasi.

### **BAB V PENUTUP**

Bab ini berisi kesimpulan yang merangkum hasil dari penelitian ini, serta saransaran yang dapat diberikan untuk penelitian lanjutan. Kesimpulan diambil berdasarkan hasil dan pembahasan yang telah dilakukan, dan saran diberikan terkait pengembangan lebih lanjut dari metode atau pendekatan yang digunakan.