

BAB I

PENDAHULUAN

1.1 Latar Belakang

Hati, atau *Liver*, memiliki kemampuan unik untuk regenerasi, yaitu menggantikan sel-sel yang rusak. Organ ini memiliki peran krusial dalam berbagai proses metabolisme tubuh. Dalam metabolisme lemak, hati menghasilkan empedu dan kolesterol yang berperan dalam pencernaan lemak. Selain itu, hati juga berperan dalam metabolisme protein, menghasilkan asam amino yang digunakan untuk membangun protein yang dibutuhkan dalam melawan infeksi dan menetralkan amonia [1].

Penyakit *Liver* merupakan masalah kesehatan utama bagi jutaan orang di seluruh dunia, prediksi dan identifikasi dini sangat penting untuk mengambil tindakan yang tepat pada tahap awal penyakit ini [2]. Penyakit hati atau *Liver* yang sudah akut sangat mempengaruhi fungsi-fungsi hati, penyakit hati dapat diketahui dari munculnya gejala klinis maupun fisik yang timbul pada pasien [3]. Penyebab penyakit hati bisa mencakup infeksi, cedera, pengaruh obat-obatan atau zat berbahaya, dan keabnormalan genetik. Dampak yang diberikan oleh penyakit *Liver* ini berbeda-beda tergantung pada tingkat keparahan dan respons pengobatan yang dilakukan oleh individu. Semakin parah penyakit *Liver* maka akan memberikan dampak yang lebih serius bahkan bisa menyebabkan kematian. Respon pengobatan pada penyakit *Liver* juga sangat berpengaruh pada akibat yang ditimbulkannya pada kesehatan pasien [4]. Kronisnya penyakit pada hati merupakan suatu permasalahan dalam kesehatan yang disebabkan oleh konsumsi alkohol memiliki permasalahan yang signifikan di seluruh dunia. Kejadian kematian sekitar 3,3 juta atau 5,9 % dari seluruh kematian dari jumlah yang mengkonsumsi alkohol setiap tahunnya,

serta kondisi kematian yang disebabkan oleh penyakit hati karena alkohol, hal tersebut berdasarkan World Health Organization (WHO), 2019 [5].

Menurut Organisasi Kesehatan Dunia (WHO), penyakit hati, terutama hepatitis, masih menjadi masalah kesehatan global yang signifikan. Secara global, diperkirakan lebih dari 354 juta individu hidup dengan hepatitis B dan C, yang merupakan penyebab utama sirosis hati dan kanker hati. Kedua kondisi ini berkontribusi secara signifikan terhadap beban penyakit hati di seluruh dunia. Di Indonesia, hepatitis B dan C merupakan penyebab utama sebagian besar kasus sirosis dan kanker hati, yang mengakibatkan lebih dari 51.000 kematian setiap tahunnya. Diperkirakan sekitar 18 juta orang di Indonesia terinfeksi hepatitis B, dan hampir setengah dari mereka berisiko mengalami kondisi kronis seperti sirosis atau kanker hati.

Kesulitan dalam mengidentifikasi gejala penyakit hati sejak dini menjadi kendala signifikan. Deteksi dini gejala penyakit hati sangat penting untuk meningkatkan peluang pasien bertahan hidup. Penyakit hati memerlukan penanganan sedini mungkin untuk mencegah komplikasi penyakit lain yang dapat timbul akibat gangguan atau kerusakan hati [4]. Penanganan penyakit hati yang lebih awal berkorelasi dengan penurunan risiko kondisi buruk yang diakibatkan oleh penyakit ini. Namun, diagnosis penyakit hati seringkali menantang karena gejalanya yang tidak spesifik dan sulit dideteksi secara dini. Oleh karena itu, pengembangan sistem diagnosis penyakit hati menjadi relevan dan bermanfaat. Hasil dari sistem diagnosis yang dikembangkan dapat membantu dokter dan tenaga medis dalam membuat keputusan terkait penanganan medis pasien[4].

Dalam bidang ilmu komputer, data mining adalah metode yang digunakan untuk menganalisis kumpulan data. Machine learning, sebagai bagian dari data mining, adalah teknik pembelajaran mesin yang berfokus pada identifikasi pola berdasarkan model untuk klasifikasi dan prediksi data baru.

Secara umum, machine learning melibatkan empat tahapan: (1) definisi masalah, (2) pengumpulan dan persiapan data, (3) pembuatan model, dan (4) prediksi model. Machine learning dapat diterapkan untuk menyelesaikan masalah sesuai dengan kebutuhan masing-masing bidang [3]. Dalam konteks medis, dengan memanfaatkan data klinis dan kesehatan pasien, model prediktif berbasis machine learning memungkinkan identifikasi individu yang berisiko, sehingga intervensi medis dapat dilakukan lebih awal dan tepat sasaran. Penelitian ini secara khusus mengevaluasi tiga algoritma yang umum digunakan dalam klasifikasi data kesehatan, yaitu *Random Forest*, *C5.0*, dan *Support Vector Machine* (SVM). *Random Forest* dikenal karena akurasi yang tinggi karena menggabungkan banyak pohon keputusan, sehingga hasilnya lebih stabil dan mengurangi risiko overfitting. *C5.0*, sebagai pengembangan dari *C4.5*, mampu menghasilkan pohon keputusan yang lebih sederhana dan efisien dalam menangani dataset yang kompleks. Sementara itu, *Support Vector Machine* efektif untuk data berdimensi tinggi dan dikenal karena kemampuannya untuk memisahkan kelas secara optimal.

Penelitian ini bertujuan untuk mengevaluasi dan membandingkan kinerja algoritma *Random Forest*, *C5.0*, dan *Support Vector Machine* (SVM) dalam memprediksi risiko penyakit hati. Melalui pengukuran metrik akurasi, presisi, dan *Recall*, penelitian ini diharapkan dapat memberikan rekomendasi algoritma yang paling efektif dalam mendeteksi risiko penyakit hati. Temuan ini diharapkan dapat menjadi acuan bagi para profesional kesehatan dan pembuat kebijakan dalam memilih metode yang tepat untuk mendukung upaya pencegahan dan pengobatan penyakit hati.

1.2 Ruang Lingkup Penelitian

- a. Data yang diprediksi merupakan data *Liver* pada balita dengan jumlah 30692 set data yang diambil dari *Kaggle*.
- b. Menganalisis dan membandingkan Tingkat akurasi dari algoritma *Random Forest*, *C 5.0*, dan *Support Vector Machine* dalam memprediksi risiko penyakit *Liver*.
- c. Evaluasi algoritma akan menggunakan metrik evaluasi seperti akurasi, *precision*, dan *Recall* untuk membandingkan performa masing-masing.

1.3 Rumusan Masalah

Berdasarkan masalah dan alternatif solusi yang diusulkan, maka ditetapkan rumusan masalah sebagai berikut. “Bagaimana performa algoritma *Random Forest*, *C 5.0* dan *Support Vector Machine* dalam memprediksi penyakit *Liver* metrik akurasi, akurasi, *precision*, dan *Recall*, serta algoritma mana yang memiliki performa terbaik dalam memprediksi risiko penyakit *Liver*?”.

1.4 Tujuan Penelitian

Tujuan pada penelitian ini adalah sebagai berikut :

- a. Menganalisis performa algoritma *Random Forest*, *C 5.0* dan *Support Vector Machine* dalam memprediksi penyakit *Liver* berdasarkan metrik akurasi, akurasi, *precision*, dan *Recall*. Membandingkan ketiga algoritma tersebut untuk menentukan algoritma dengan performa terbaik dalam memprediksi risiko penyakit *Liver*.

1.5 Manfaat Penelitian

Manfaat dari hasil penelitian ini adalah sebagai berikut :

- a. Memberikan kontribusi dalam bidang ilmu pengetahuan dengan menambah wawasan tentang efektivitas berbagai algoritma klasifikasi dalam mendeteksi risiko kesehatan, khususnya penyakit *Liver*.
- b. Memberikan rekomendasi algoritma yang paling efektif untuk mendukung upaya penanganan penyakit *Liver* di Indonesia.
- c. Meningkatnya efisiensi proses deteksi dini penyakit *Liver* secara lebih cepat.

1.6 Sistematika Penulisan

Penulisan proposal ini dibagi dalam 3 bab, dengan sistematika sebagai berikut.

a. Bab I Pendahuluan

Bab ini berisi antara lain latar belakang, ruang lingkup penelitian, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

b. Bab II Tinjauan Pustaka

Pada bab ini dijelaskan tentang penelitian yang terkait dan teori-teori yang digunakan penelitian ini.

c. Bab III Metode Penelitian

Bab ini membahas tentang objek penelitian dan metode yang digunakan dalam menyelesaikan permasalahan.

d. Bab IV Hasil dan Pembahasan

Bab ini menjelaskan hasil penelitian dan pembahasan terkait penelitian.

e. Bab V Kesimpulan dan Saran

Bab ini menyimpulkan hasil dan penelitian guna mendapatkan hasil dari prediksi *Liver* pada balita menggunakan metode Decision Tree yang dijadikan sebagai jawaban dari rumusan masalah yang telah disampaikan sebelumnya.