

# 1 PENDAHULUAN

## 1.1 Latar Belakang

Seiring dengan perkembangan zaman dan teknologi, ukuran *Database* saat ini sudah berkembang menjadi sangat besar ke dalam ukuran *terabyte*. Di dalam suatu tumpukan data pada *database* tersebut mungkin terdapat informasi-informasi tersembunyi yang penting atau akan menjadi penting pada saat-saat tertentu. Akan tetapi untuk menemukan hal tersebut sama saja seperti mencari jarum dalam tumpukan jerami. Oleh karena itu dapat disimpulkan bahwa untuk mengubah data mentah untuk dapat menjadi suatu informasi yang berguna bukan merupakan hal yang mudah.

Data mentah (*raw data*) sering dianggap tidak berguna dikarenakan ukurannya yang begitu besar sehingga sulit dianalisa. Oleh karena itu perlu dilakukan pengenalan pola dari data mentah tersebut. Solusinya adalah dengan *data mining*. Banyak organisasi di dunia telah menggunakan data mining untuk mencari dan menarik kesimpulan dari data yang mereka miliki.

Penggunaan *data mining* dilatar belakangi oleh banyaknya data yang diperoleh oleh berbagai institusi, perusahaan atau organisasi. Berlimpahnya data ini merupakan akumulasi data transaksi yang terekam selama bertahun-tahun. Klasifikasi merupakan salah satu fungsi yang umum diterapkan dalam *data mining*.

Masalah klasifikasi saat ini banyak sekali ditemukan dalam kehidupan sehari-hari misalnya dalam bidang manajemen, pemasaran, penjualan dan jadwal distribusi diantara banyak *outlet*, dalam bidang *broadcasting*, untuk memprediksi program terbaik apa yang ditayangkan selama waktu tayang utama, dan terakhir sering dipakai yaitu dalam bidang kedokteran, misalnya untuk mendiagnosis penyakit. Salah satu masalah yang sering terjadi yaitu dalam mendiagnosis penyakit typhus dan demam berdarah, karena keduanya memiliki ciri-ciri dan gejala yang hampir sama sehingga sulit membedakannya, namun sebenarnya berbeda. Pada penelitian ini digunakan data *medical record* pasien suatu rumah sakit untuk penyakit demam berdarah dan typhus yang dibuat suatu modelan agar dapat diketahui faktor apa saja yang berpengaruh dalam mendiagnosis penyakit tersebut.

*Problem* seperti ini dalam *data mining* termasuk dalam jenis klasifikasi. Pada tugas akhir ini dibuat suatu model *decision tree* untuk memprediksi seorang pasien masuk dalam kategori typhus, DBD atau komplikasi keduanya dilihat dari data *medical record* pasien pada suatu rumah sakit berdasarkan ciri-ciri dan gejala pasien. Untuk membantu proses klasifikasi *data*, penulis menggunakan Algoritma yang sering digunakan untuk proses klasifikasi data dengan nilai akurasi yang tinggi yaitu Algoritma C5.0. Pemilihan algoritma ini dikarenakan Algoritma C5.0 tidak menggunakan vektor jarak untuk mengklasifikasikan obyek sehingga cocok untuk *data* observasi dengan atribut yang bersifat *numerik* ataupun atribut yang bernilai nominal yaitu bersifat *kategorik* dimana tiap nilainya tidak bisa dijumlahkan atau dikurangkan contohnya seperti bentuk, warna dan rasa. Selain itu algoritma ini juga menghasilkan *tree* dengan jumlah cabang per *node* bervariasi sehingga dapat membentuk pohon keputusan yang lebih efisien.

## 1.2 Rumusan Masalah

Berdasarkan uraian dari latar belakang yang telah dikemukakan di atas, permasalahan yang akan dibahas yaitu bagaimana mengklasifikasikan dan memprediksi penyakit Demam Berdarah dan Tipes yang diderita oleh pasien agar mendapatkan hasil dengan tingkat akurasi yang tinggi?

## 1.3 Tujuan

Penelitian ini bertujuan untuk membentuk *rule* guna mengklasifikasikan prediksi penyakit typhus atau DBD yang diderita oleh pasien berdasarkan ciri-ciri dan gejalanya dengan menggunakan teknik klasifikasi pohon keputusan dan algoritma C5.0 untuk menghasilkan rule yang seakurat mungkin.

## 1.4 Batasan Masalah

Untuk lebih memfokuskan penelitian dan menyederhanakan permasalahan, peneliti menentukan batasan masalah penelitian. Batasan yang digunakan dalam penelitian ini adalah:

1. Model dikembangkan untuk kasus klasifikasi tiga kelas yaitu typhus, DBD atau komplikasi keduanya.
2. Data yang digunakan untuk pengujian model adalah *dataset* kasus nyata yang didapatkan dari *medical record* pasien rumah sakit Sartika Asih.
3. Penelitian dilakukan dengan menggunakan metode pohon keputusan dan algoritma C5.0.

## 1.5 Metodologi Penyelesaian Masalah

Metodologi yang digunakan dalam menyelesaikan Tugas Akhir ini adalah sebagai berikut :

- a) Melakukan studi literatur mengenai algoritma C5.0
- b) Mengidentifikasi masalah ciri penyakit Demam Berdarah dan Typhus.
- c) Pengambilan data dari rumah sakit.
- d) Melakukan perancangan dan permodelan pada sistem yang akan diuji.
- e) Melakukan pengujian perangkat lunak dengan melihat performansi dan akurasi algoritma C5.0 terhadap klasifikasi kedua penyakit tersebut.
- f) *Testing* dan Analisis Hasil.
  - a. *Testing*  
Melakukan pengujian sistem yang telah dibangun menggunakan data uji dengan berbagai skenario pengujian.
  - b. Analisis Hasil  
Untuk membuktikan bahwa tingkat performansi sistem yang dibangun telah mencapai tingkat yang baik, maka dilakukan analisis terhadap performansi sistem sesuai dengan skenario pengujian yang dilakukan.
- g) Penyusunan laporan  
Menyusun laporan dan dokumentasi terhadap penelitian yang telah dilakukan, serta membuat kesimpulan dari hasil analisis tersebut dengan mengikuti

ketentuan yang telah ditetapkan oleh institusi. Laporan yang dibuat antara lain adalah buku Tugas Akhir dan jurnal.

## **1.6 Sistematika Penulisan**

Dalam penulisan Tugas Akhir ini dibagi menjadi beberapa bab yang meliputi hal-hal sebagai berikut :

1. Bab I Pendahuluan  
Berisi latar belakang, tujuan, rumusan masalah, batasan masalah, tahapan penyelesaian masalah yang digunakan, serta sistematika penulisan yang memuat susunan penulisan Tugas Akhir ini.
2. Bab II Dasar Teori  
Bab ini membahas tentang teori-teori yang mendukung yaitu, informasi mengenai penyakit demam berdarah dan typhus serta algoritma C5.0
3. Bab III Perancangan Sistem  
Bab ini membahas tentang perancangan sistem untuk membuat sistem prediksi penyakit dengan algoritma C5.0.
4. Bab IV Implementasi dan Analisa Sistem  
Bab ini membahas tentang pengujian sistem & analisis terhadap implementasi Algoritma C5 dalam memprediksi penyakit demam berdarah dan typhus
5. Bab V Kesimpulan dan Saran  
Bab terakhir ini memberikan kesimpulan hasil penelitian pada Tugas Akhir yang telah dilakukan dan saran terhadap pengembangan ke depan.