

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Kesehatan merupakan peranan yang sangat penting untuk mempertahankan kehidupan manusia, dengan kesehatan yang baik manusia dapat melakukan kegiatan produktif dalam sosialisasi atau ekonomi untuk mencapai tujuan hidup mereka. Salah satu penyakit yang dapat menyebabkan komplikasi dan kematian adalah *diabetes*. *Diabetes* bukan hanya penyebab utama kematian dini di dunia, penyakit ini dapat menyebabkan kebutaan, gagal ginjal, dan bahkan penyakit jantung [1]. *WHO* memperkirakan jumlah penderita *diabetes* di Indonesia akan meningkat dari 8,4 juta pada tahun 2000 menjadi sekitar 21,3 juta pada tahun 2030, dan *International Diabetes Federation (IDF)* juga memperkirakan jumlah penderita *diabetes* pada tahun 2009 akan mencapai 7,0 juta menjadi 12,0 juta pada tahun 2030 [2].

Diabetes adalah kondisi ketika kandungan gula atau kadar glukosa dalam darah melebihi normal yang disebabkan oleh pankreas tidak memproduksi cukup insulin untuk tubuh atau tubuh tidak menggunakan insulin secara efektif. Insulin adalah hormon yang mengontrol gula darah, keseimbangan gula darah yang berasal dari makanan yang dimakan orang tersebar ke sel darah dalam tubuh yang dapat menghasilkan energi [3]. Penderita *diabetes* terdapat faktor resiko yang diperkirakan seperti penderita melakukan pola hidup tidak sehat seperti konsumsi

alkohol serta kebiasaan merokok atau perubahan gaya hidup contohnya pola makan yang tidak sehat dan tidak seimbang, pasien juga kurang melakukan aktivitas fisik yang menyebabkan kelebihan berat badan (*obesitas*) dan tekanan darah tinggi. Faktor-faktor lain dari penderita *diabetes* seperti usia, jenis kelamin maupun keturunan. Pendiagnosis pada penderita dengan cara melakukan pengajuan pertanyaan seperti keluhan yang dirasakan penderita dengan mendapatkan hasil kesimpulan berupa tingkat resiko penyakit *diabetes*.

Di bidang medis terdapat banyak catatan pasien salah satunya adalah data penyakit diabetes. Namun, jumlah data tidak dapat digunakan dengan baik tanpa informasi dan kesimpulan dari data. Seperti sulit memprediksi apakah pasien mengidap diabetes atau tidak. Sehingga diperlukan proses ekstraksi untuk mencari informasi pada data yang sebelumnya tidak diketahui yang dikenal dengan istilah *data mining*.

Data mining adalah proses menganalisis data dari berbagai sudut dan kumpulan data yang sebelumnya tidak digunakan untuk mendapatkan pengetahuan baru dengan menemukan pola tersembunyi dalam data kemudian mengubahnya menjadi informasi yang berguna [4]. Fungsi utama dari *data mining* adalah klasifikasi. Salah satu algoritma untuk klasifikasi yaitu dengan menggunakan algoritma C4.5. Algoritma C4.5 adalah metode yang digunakan untuk klasifikasi yang menghasilkan model berupa pohon keputusan yang digunakan untuk mengklasifikasi [5].

Dalam mendiagnosis diperlukan suatu metode untuk memprediksi *diabetes* secara lebih akurat dan efektif. Oleh karena itu peneliti menggunakan algoritma C4.5 untuk memprediksikan penyakit *diabetes* sehingga memberikan hasil yang akurat dari proses evaluasi. Hal inilah yang melatar belakangi penulis untuk mengambil sebuah penelitian dengan judul **“PENERAPAN DATA MINING UNTUK PREDIKSI PENYAKIT DIABETES MENGGUNAKAN ALGORITMA C4.5 ”**.

1.2 RUMUSAN MASALAH

Berdasarkan latar belakang peneliti tersebut, maka rumusan masalah penelitian ini adalah “Bagaimana Menerapkan *Data Mining* Untuk Prediksi Penyakit *Diabetes* Menggunakan *Algoritma C4.5*?”

1.3 BATASAN MASALAH

Agar pembahasan ini tidak menyimpang dari apa yang telah dirumuskan, maka dibutuhkan batasan-batasan. Batasan-batasan dalam penelitian ini adalah sebagai berikut:

1. Pada penelitian ini peneliti hanya menggunakan metode prediksi dengan algoritma C4.5.
2. Data yang digunakan merupakan data yang diambil dari dataset *Kaggle*. Atribut yang di gunakan pada penelitian ini yaitu : Jenis Kelamin, Umur, *Polyuria, Polydipsia, Sudden weight lost, Weakness, Polyphagia, Genital thrush, Visual blurring, Itching, Irritability, Delayed healing, Partial paresis, Muscle stiffness, Alopecia, Obesity, Class*.
3. Alat bantu (*tools*) yang digunakan penelitian ini adalah *Rapid Miner*.

1.4 TUJUAN PENELITIAN

Adapun tujuan yang ingin dicapai dalam penelitian ini adalah :

1. Menerapkan teknik *Data Mining* dengan algoritma C4.5 dalam memprediksi penyakit *Diabetes*.
2. Mengevaluasi hasil perhitungan algoritma C4.5 pada penderita penyakit *Diabetes*.

1.5 MANFAAT PENELITIAN

Adapun manfaat dari penelitian tugas akhir yang dilakukan adalah :

1. Dapat mengetahui status penderita *Diabetes* dengan melihat tingkat akurasi yang tinggi.
2. Dapat digunakan sebagai acuan untuk penelitian selanjutnya.

1.6 SISTEMATIKA PENULISAN

Untuk memberikan gambaran umum mengenai keseluruhan penulisan ilmiah, dapat dilihat melalui sistematika penulisan yang meliputi :

BAB I : PENDAHULUAN

Bab ini menguraikan tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan.

BAB II : LANDASAN TEORI

Pada bab landasan teori ini membahas teori-teori dan pendapat para ahli yang berhubungan dengan pembahasan yang dianalisis. Teori-teori yang digunakan antar lain mengenai *Data Mining*, *Diabetes*, C4.5, dan *Rapid Miner*.

BAB III : METODOLOGI PENELITIAN

Dalam bab ini berisi metode pengumpulan data, prosedur penelitian, dan metode analisis berupa pendekatan penyelesaian masalah yang dilakukan untuk mendukung penelitian.

BAB IV : ANALISIS

Bab ini berisi tentang objek penelitian, alat dan bahan, hasil, implementasi, analisis dan pembahasan penelitian yang berupa pengujian mengenai algoritma yang digunakan untuk memprediksi penyakit *Diabetes*.

BAB V : PENUTUP

Bab ini disajikan simpulan dari hasil penelitian dan juga saran-saran yang ditujukan kepada semua pihak yang bersangkutan.