

BAB I

PENDAHULUAN

1.1. Latar Belakang

Kesehatan yaitu peranannya sangat penting untuk menunjang kehidupan manusia, dengan memiliki kesehatan yang baik, manusia dapat melakukan aktifitas dengan produktif dalam sosialisasi atau ekonomi untuk mencapai tujuan hidup. Salah satu penyakit yang dapat mengakibatkan komplikasi bahkan kematian adalah penyakit diabetes. Diabetes bukan hanya penyebab dari kematian prematur di dunia, penyakit ini bisa menyebabkan kebutaan, gagal ginjal, dan bisa juga menyebabkan penyakit jantung. (Infodatin, 2020).

Diabetes Miletus merupakan suatu kondisi kronis bisa berlangsung seumur hidup dan akan mempengaruhi kemampuan tubuh dalam menggunakan energi makanan yang telah dicerna. (Suryati et al., 2020). Menurut *International Diabetes Federetion* (IDF) kusus pasien pengidap diabetes miletus di dunia meningkat setiap tahunnya. Tahun 2011 terdapat 366 juta jiwa, tahun 2013 terdapat 382 juta jiwa, tahun 2015 yaitu 415 juta jiwa, sedangkan tahun 2017 naik hingga 425 juta jiwa, dan pada tahun 2019 terdapat 463 juta jiwa. Diperkirakan 2045 pengidap penyakit diabetes terus mengalami kenaikan menjadi 700 juta jiwa yang terdiagnosa (Alisa et al, 2020).

Berdasarkan WHO atau organisasi kesehatan dunia memprediksi pengidap penyakit diabetes miletus di Republik Indonesia dari 8,4 juta jiwa di tahun 2020, dan pada tahun 2030 naik sampai 21,3 juta jiwa. Sedangkan menurut *International Diabetes Federation* (IDF) juga pada tahun 2009 pengidap penyakit diabetes 7,0 juta jiwa sampai 12,0 juta jiwa tahun 2030. Pada tahun 2030 menurut prediksi WHO dan IDF penderita penyakit diabetes miletus di Indonesia naik 2-3 kali lipat (Ente et al., 2020).

Pendekatan data mining menjadi sangat penting dalam bidang Kesehatan untuk mengambil keputusan berdasarkan data klinis yang besar (Fiandra et al., 2017). Data mining ialah proses untuk menemukan pola atau informasi dari pengolahan data yang belum ditemukan dalam data yang dipilih dengan menggunakan cara atau teknik tertentu sesuai dengan data. (Mardi, 2017). Dengan diterapkannya data mining diharapkan dapat

menjadi suatu informasi untuk penyakit diabetes di Indonesia ataupun dunia, sehingga angka penderita penyakit diabetes dapat menurun.

Teknik klasifikasi termasuk kedalam bagian metode *supervised learning* yaitu diperlukannya data latihan dalam membangun pola untuk model klasifikasinya. Teknik klasifikasi memiliki beberapa pilihan algoritma, algoritma yang masuk kedalam metode *supervised learning* yaitu *k-Nearest-Neighbor* (k-NN), Naïve Bayes, *Support Vector Machine*, ID3, dan algoritma C4.5 (Khotimah & Istiawan, 2018).

Algoritma C4.5 termasuk dalam algoritma klasifikasi yang menghasilkan pohon keputusan dan bisa diolah dengan data diskrit dan numerik, selain itu algoritma C4.5 dapat menghasilkan cara yang mudah untuk diinterpretasikan. C4.5 telah dicoba oleh diberbagai kasus klasifikasi seperti bidang kepegawaian, perdagangan, medis dan bidang lainnya (Yuningsih et al., 2020).

Pada penelitian Rahman analisa pasien penyakit liver menggunakan metode Naïve Bayes dan *decision tree* C4.5 menyimpulkan metode *decision tree* C4.5 memiliki akurasi yang paling besar 70.29% sedangkan algoritma Naïve Bayes menghasilkan akurasi 67.05% (Rahman, 2020). Penelitian lain oleh Leidiyana & Permana untuk meningkatkan kualitas pemilihan calon karyawan C4.5 menggunakan *confusion matrix* mendapatkan hasil akurasi sebesar 87.5% (Leidiyana & Permana, 2020). Penelitian lainnya yang menggunakan perbandingan C4.5 dan Naïve Bayes untuk menentukan dosen tetap yang dilakukan oleh Sadikin, C4.5 menghasilkan akurasi tertinggi 91.89% sementara itu Naïve Bayes menghasilkan nilai akurasi sebesar 83.78% (Sadikin et al., 2020).

Berdasarkan permasalahan di diatas maka penulis tertarik untuk melakukan penelitian yang berjudul “Klasifikasi Penderita Penyakit Diabetes Menggunakan Algoritma C4.5”

1.2. Rumusan Masalah

Dari latar belakang diatas yang telah disimpulkan, penulis merumuskan beberapa masalah dari latar belakang sebagai berikut:

1. Bagaimana cara menerapkan C4.5 dalam mengklasifikasikan penderita penyakit diabetes?
2. Berapakah hasil akurasi dari C4.5 dalam mengklasifikasikan penderita penyakit diabetes?

1.3. Tujuan Penelitian

Tujuan penelitian ini berdasarkan dari rumusan masalah diatas, dengan itu penulis merumuskan tujuan penelitian sebagai berikut:

1. Menerapkan algoritma C4.5 dalam mengklasifikasikan penderita penyakit diabetes.
2. Mengetahui akurasi dari hasil menerapkan algoritma C4.5 dalam mengklasifikasikan penderita penyakit diabetes.

1.4. Manfaat

Manfaat dari penelitian sebagai berikut:

1. Menjadikan informasi tentang algoritma klasifikasi C4.5 dalam mengklasifikasikan penderita penyakit diabetes.
2. Memberikan informasi tentang nilai akurasi dari algoritma klasifikasi C4.5 dalam mengklasifikasikan penderita penyakit diabetes.

Dapat dijadikan sumber refensi untuk penelitian berikutnya