

# **BAB I**

## **PENDAHULUAN**

### **1.1 LATAR BELAKANG**

Pada era saat ini seiring dengan berkembangnya teknologi computer yang semakin pesat segala bidang pekerjaan mulai menggunakan penerapan teknologi, maka perkembangan teknologi juga berdampak pada meningkatnya kebutuhan akan perkembangan dunia, sehingga mampu mengubah metode dalam melakukan pengolahan data. Disuatu sisi perkembangan ilmu pengetahuan dan teknologi akan mempermudah persaingan local maupun global. Untuk menghadapi persoalan tersebut.

Pada saat ini penyakit Penyakit ginjal kronis dapat berasal dari berbagai penyakit serta masalah kesehatan lain yang memberikan tekanan pada ginjal. Yang dimana suatu kondisi yang dapat dikenal sebagai penyakit ginjal kronik (Chronic Kidney Disease) ditandai dengan penurunan fungsi ginjal yang menetap (kronis) dan cukup parah yang disebabkan oleh beberapa penyakit ginjal. Penyakit ini dapat berkembang dan biasanya tidak ada obatnya. Gejala khas penyakit ini antara lain ketidakmampuan makan, mual, muntah, sakit kepala ringan, sesak napas, kelelahan, edema lengan dan kaki, serta uremia [1].

Gejala pada penderita penyakit ginjal kronik biasanya tidak begitu terlihat. Umumnya, gejala penyakit ginjal kronik baru terasa pada saat terjadinya gangguan metabolisme tubuh yang dimana ginjal tidak dapat menyaring jumlah racun yang banyak. Berdasarkan diagnosa medis, rata-rata kejadian penyakit ginjal kronik

(PGK) yang ada di Indonesia pada tahun 2013 sebanyak 0,2%, dengan frekuensi tertinggi dan terbanyak di Sulawesi Tengah sebanyak 0,5%, disusul aceh, gorontalo, dan sulawesi utara sebesar 0,4%. Prevalensi penyakit ginjal kronik sebanyak 0,2% yang ada di wilayah Sumatera Utara, 0,3% di nusa Tenggara timur, Sulawesi Selatan, lampung, jawa barat, jawa Tengah, Yogyakarta, dan jawa timur, serta 0,3% di Nusa Tenggara Timur [2]. Pasien mengalami gagal ginjal stadium 5, yang ditentukan oleh laju filtrasi glomerulus (GFR), sebelum timbulnya penyakit ginjal stadium akhir. Di stadium akhir fungsi ginjal akan mengalami penumpukan cairan yang harusnya dapat dikeluarkan dari urine. Maka dari itu kenapa penyakit ginjal kronik harus dilakukan klasifikasi, karena mengklasifikasi penyakit ginjal kronik untuk mempermudah dalam mengenali penyakit ginjal kronik, membandingkan, dan mempelajari penyakit ginjal kronik.

Namun, untuk meringkas data secara otomatis dan mengekstrak informasi atau pola tersembunyi dari data, yang semakin rumit, diperlukan alat yang sesuai. Diperlukan suatu alat yang dapat mengevaluasi data dan mengekstrak informasi yang dapat membantu dalam proses pengambilan keputusan, karena semakin banyaknya data yang disimpan dalam file, database, dan gudang data. Penambangan data adalah istilah yang digunakan untuk alat ini. Penambangan data adalah proses penggunaan alat atau pendekatan khusus untuk mencari pola atau informasi menarik dalam data yang dipilih. Ada banyak pendekatan, metode, atau algoritma penambang data yang berbeda yang berbeda. Tujuan dan proses penemuan pengetahuan basis data (KDD) secara keseluruhan memiliki dampak yang signifikan terhadap pemilihan metode atau algoritma terbaik [2].

Metode-metode yang terdapat di dalam data mining yang dimana digunakan untuk memprediksi salah pendekatan Naive Bayes yang eksklusif. Penggunaan teorema Bayes dengan asumsi independensi yang sangat kuat merupakan dasar dari metode peramalan langsung berbasis probabilitas yang dikenal dengan nama Naive Bayes. Pengklasifikasi probabilistik mudah yang disebut Dengan menambahkan distribusi probabilitas dan kumpulan nilai dari kumpulan data tertentu, algoritma Naive Bayes menentukan kumpulan probabilitas. Algoritme ini menggunakan teorema Bayes dan mengasumsikan bahwa semua fitur bersifat independen atau tidak terkait berdasarkan nilai variabel kelas [3].

Pendekatan ini akan menghasilkan konstruksi pohon keputusan. Metode C4.5 adalah salah satu pendekatan penambangan data decision tree terkenal yang dimana dapat dipakai untuk melakukan prediksi. Dimana metode C4.5 ialah sebuah sekelompok algoritma untuk masalah klasifikasi data mining dan machine learning. Algoritma ID3 telah berkembang menjadi algoritma C4.5. J. Rose Quinlan, seorang peneliti kecerdasan buatan, mengembangkan algoritma C4.5 dan ID3 pada akhir tahun 1970-an. Atribut teratas merupakan root dan atribut bawah sebagai daun pada pohon keputusan atau decision tree yang dibuat oleh algoritma C4.5. [4].

Komparasi ini memiliki tujuan untuk menentukan metode mana yang memiliki akurasi yang tinggi dan akurat dan optimal dalam melakukan klasifikasi penyakit ginjal kronik. Berdasarkan uraian latar belakang, maka dari itu diusukan sebuah penelitian dengan judul **“KOMPARASI METODE NAÏVE BAYES DAN C4.5 UNTUK KLASIFIKASI PENYAKIT GINJAL KRONIK”**

## 1.2 RUMUSAN MASALAH

Adapun rumusan masalah yang ada dalam pada penelitian ini yaitu:

1. Bagaimana penerapan metode Naïve Bayes Dan C4.5 untuk klasifikasi penyakit ginjal kronik?
2. Bagaimana hasil komparasi metode Naïve Bayes dan C4.5 untuk klasifikasi penyakit ginjal kronik?

## 1.3 BATASAN MASALAH

Penulis mengusulkan batasan masalah berikut agar kita dapat tetap pada topik dan menghindari kesalahan dengan mengidentifikasi penyakit ginjal kronis sebagai sesuatu yang di luar topik.:

1. Dataset tersebut didapatkan dari website Kaggle.com.
2. Penelitian ini digunakan untuk komparasi metode Naïve Bayes dan C4.5 untuk klasifikasi penyakit ginjal kronik.
3. Metode yang digunakan Naïve Bayes Dan C4.5.
4. Alat bantu analisa menggunakan WEKA dan Microsoft excel

## 1.4 TUJUAN PENELITIAN

Tujuan akhir dari penelitian ini akan dijelaskan lebih rinci di bawah ini, sebagai berikut:

1. Melakukan komparasi metode Naïve Bayes dan C4.5 untuk klasifikasi penyakit ginjal kronik

2. Mendapatkan nilai akurasi yang akurat dan optimal pada metode Naïve Bayes dan C4.5

#### 1.5 MANFAAT PENELITIAN

Manfaat dari penelitian ini akan dijelaskan lebih rinci di bawah ini, sebagai berikut :

1. Memudahkan dalam melakukan klasifikasi penyakit ginjal kronik
2. Menghasilkan nilai akurasi yang akurat pada metode Naïve Bayes Dan C4.5
3. Pembaca dapat menggunakannya sebagai pedoman dalam melakukan kajian dan menciptakan suatu sistem yang jauh lebih unggul.

#### 1.6 SISTEMATIKA PENULISAN

Hal ini dapat dilihat melalui sistematika penulisan yang meliputi, memberikan gambaran umum terhadap seluruh penulisan ilmiah :

#### BAB I : PENDAHULUAN

Bagian pertama mengkaji definisi isu dan konteks sejarah serta tujuan dan keuntungan dari proses penelitian dan penulisan.

#### BAB II : LANDASAN TEORI

Teori dan pendapat para ahli terkait dengan topik yang dianalisis tercakup dalam bab landasan teori. Penambangan data, anemia, naif bayes, dan WEKA adalah beberapa teori yang diterapkan.

#### BAB III : METODOLOGI PENELITIAN

Metode yang digunakan yaitu pengumpulan data, prosedur penelitian, dan teknik analisis strategi pemecahan masalah semuanya tercakup dalam bab ini.

#### BAB IV : ANALISIS DAN HASIL

Pada bagian ini merinci mengenai penelitian tentang cara menguji algoritma yang digunakan untuk mengidentifikasi pasien penyakit ginjal kronik. Di dalamnya juga memuat objek penelitian, alat dan bahan, hasil, pelaksanaan, analisis, dan pembahasan hasil.

#### BAB V : PENUTUP

Demi kepentingan semua pihak yang berkepentingan, bab ini merangkum kesimpulan dan rekomendasi temuan.