

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Ginjal adalah salah satu organ vital yang berperan penting dalam menjaga kestabilan tubuh, ginjal mengatur keseimbangan cairan tubuh, elektrolit dan asam basa dengan cara menyaring darah melalui ginjal, reabsorpsi selektif air, serta mengekresi kelebihanannya sebagai kemih [1].

Menurut studi Global Burde of Disease, penyakit ginjal kronik menduduki peringkat ke-27 dalam daftar penyebab jumlah total kematian di seluruh dunia pada tahun 1990, tetapi naik menjadi ke-18 pada tahun 2010 [2].

Menurut buku panduan evaluasi profesi yang disusun oleh Ns. Henni Kusuma , M.Kep., Sp.Kep.MB., dkk menyatakan bahwa factor penyebab penyakit ginjal kronik yang umum adalah Hipertensi, Diabetes, Umur, Riwayat keluarga penderita diabetes, hipertensi, dan penyakit ginjal kronik.[3]

Pada gagal ginjal kronis, fungsi ginjal menurun secara drastis yang berasal dari nefron. Inefisiensi GFR (Glomerular Filtration Rate) dari ginjal sekitar 20-50%. Pada penurunan fungsi rata-rata 50%, muncul tanda dan gejala azotemia sedang, poliuri, nokturia, hipertensi dan sesekali terjadi anemia. Selama terjadi kegagalan fungsi ginjal, keseimbangan cairan dan elektrolit pun terganggu.

Penurunan fungsi ginjal menyebabkan kadar ureum, kreatinin dan asam urat meningkat sehingga menyebabkan rasa mual dan penurunan nafsu makan.[4]

Penyakit ginjal kronik (PGK) merupakan salah satu masalah kesehatan serius yang ada di dunia. Berdasarkan data dari Riskesdas penyakit ginjal kronik naik dari 2% menjadi 3,8%. Karena terjadinya peningkatan sehingga pasien PGK yang menjalani terapi HD mengalami peningkatan sebanyak 78.281 pasien, 25.446 pasien baru dan 52.835 pasien lama yang masih aktif. Salah satu masalah yang menjadi kegagalan hemodialisa adalah kepatuhan klien. Dampak ketidakpatuhan dapat mempengaruhi penurunan kesehatan, selain itu juga bisa meningkatkan angka kematian dan kecacatan. Maka dari itu pasien HD dituntut harus patuh, salah satu faktor yang dapat mempengaruhi kepatuhan adalah selfefficacy. [5].

Mengacu pada permasalahan tersebut maka diperlukan sebuah metode yang dapat digunakan untuk mengurangi tingkat kesalahan dan kecurangan yaitu dengan menerapkan teknik data mining, Data mining ini memiliki banyak fungsi. Tetapi, yang paling utama adalah sebagai Fungsi descriptive dan fungsi predictive. Dengan memanfaatkan metode Klasifikasi ini diharapkan mampu membantu user untuk menghitung jumlah Data Penyakit Ginjal Kronik dengan skala yang besar.

Data mining adalah proses mengekstraksi untuk mendapatkan informasi penting yang sifatnya *implicit* dan belum diketahui[6].

Data mining adalah klasifikasi merupakan sebuah proses menemukan definisi kesamaan karakteristik dalam suatu kelompok atau kelas (*class*). Klasifikasi *Data mining* menjadi salah satu metode yang umum digunakan. Metode

ini dilakukan bertujuan untuk memperkirakan kelas dari suatu objek yang labelnya belum diketahui. Dalam upaya meningkatkan akurasi data penyakit ginjal kromik di indonesia menggunakan metode *Naïve Bayes*. Teknik seleksi fitur yang digunakan adalah *Information Gain*. Dengan menerapkan Teknik seleksi fitur tersebut pada data dengan menggunakan metode *Naïve Bayes* maka diharapkan nantinya dapat menghasilkan hasil penyeleksian atribut yang relevan terhadap data penyakit ginjal kromik di indonesia .

Teknik seleksi fitur *Information Gain* merupakan teknik seleksi fitur yang memakai metode scoring untuk nominal ataupun pembobotan atribut kontinu yang didiskretkan menggunakan maksimal entropy. Suatu entropy digunakan untuk mendefinisikan nilai *Information Gain*. Entropy menggambarkan banyaknya informasi yang dibutuhkan untuk mengkodekan suatu kelas.

Teknik Seleksi Fitur yang digunakan adalah fitur dengan *nilai Information Gain, Information Gain* Pada penelitian ini menggunakan algoritma naïve bayes disertai *information gain* sebagai metode seleksi fitur sebagai teknik untuk memperbaiki tingkat klasifikasi yang diterapkan untuk mengklasifikasikan Penyakit Ginjal Kronik untuk meningkatkan akurasi. [7]

Sehingga dengan menggunakan Teknik seleksi fitur *Information Gain* metode *Naïve Bayes* dalam penelitian ini untuk menghitung Klasifikasi Penyakit Ginjal Kronik mampu mengurangi fitur yang tidak relevan dan mengurangi dimensi fitur pada data sehingga akurasi yang didapat bisa maksimal.

Adapun beberapa penelitian yang terkait dari penelitian diatas yang menerapkan data mining dengan menggunakan metode klasifikasi untuk menghitung dan memprediksi klasifikasi penyakit ginjal kronik, beberapa peneliti yang melakukan penelitian tentang klasifikasi ginjal kronik: penelitian yang dilakukan oleh Andi Sri Irtawaty Tingkat akurasi pengclusteran penyakit ginjal menggunakan metode K-Means dalam penelitian ini cukup tinggi, sekitar 90% [8], lalu penelitian terkait dilakukakan oleh Arim Nur Fadillah, Jeffry Wijaya dan Riko Hidayat menggunakan penerapan algoritma *Naïve Bayes* mendapatkan hasil akurasi yang cukup tinggi sebesar 91.25% [9], lalu penelitian terkait dilalakukan oleh Robi Rianto dan Ni Made Iswari menggunakan Algoritma C4.5 dan menghasilkan akurasi yang cukup tinggi yaitu 91.50% [10].

Dengan meningkatnya jumlah pengidap penyakit gagal ginjal kronik dan perlunya diagnosis yang tepat maka perlu dilakukan penelitian untuk mencari metode yang memberikan akurasi tertinggi.

Dua penyebab utama penyakit kronis adalah diabetes dan hipertensi. Penelitian ini dimaksudkan untuk melihat tingkat faktor utama yang sering dapat menyebabkan ginjal kronik tersebut dari jumlah dataset.

Tujuan penelitian ini melakukan optimalisasi dari penelitian sebelumnya dan membandingkan hasil dari presentase penelitian sebelumnya yang dilakukakan oleh Arim Nur Fadillah, Jeffry Wijaya dan Riko Hidayat sebesar 91.50%[9] dan masih menggunakan metode yang sama yaitu metode *Naïve Bayes*. Guna meningkatkan akurasi dalam mendeteksi faktor gejala utama pada ginjal kronik

maka diterapkan dengan menggunakan seleksi fitur, Dengan menggunakan seleksi fitur dilakukan untuk mereduksi data yang ada sehingga dapat memudahkan atribut apa saja yang relevan terhadap data Penyakit ginjal kronik. Penelitian ini menggunakan seleksi fitur *Information Gain* dan Metode algoritma *Naïve Bayes*.

Metode *Naïve Bayes* merupakan sebuah pengklasifikasian probalistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang di berikan. Algoritma menggunakan *Teoroma Bayes* dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas [11]. Metode ini diharapkan membantu meningkatkan akurasi dan membantu penulis dalam memprediksi penyakit ginjal kronis.

Dari uraian yang melatar belakangi masalah diatas maka yang penulis melakukan penelitian guna memberi solusi terhadap masalah yang terjadi dengan mengangkat judul “**PENERAPAN DATA MINING UNTUK KLASIFIKASI DATA PENYAKIT GINJAL KROKIK DI INDONESIA MENGGUNAKAN METODE NAÏVE BAYES DAN SELEKSI FITUR INFORMATION GAIN**”.

1.2 RUMUSAN MASALAH

Berdasarkan latar belakang diatas, maka dapat diambil suatu rumusan masalah yaitu sebagai berikut:

1. Bagaimana menerapkan *Information Gain* dan metode *Naïve Bayes* dalam memprediksi penyakit Ginjal Kronik serta menghasilkan akurasi yang baik.

2. Bagaimana tingkat akurasi klasifikasi Penyakit Ginjal Kronik menggunakan metode *Naïve Bayes* dengan metode *Information Gain* ?

1.3 BATASAN MASALAH

Agar dalam penelitian ini dapat berjalan dengan baik dan terarah penulis menetapkan ruang lingkup penelitian meliputi:

1. Penelitian ini hanya berfokus pada dataset yang bersumber dari web kaggle.com.
2. Metode klasifikasi yang digunakan adalah *Naïve Bayes*
3. Teknik seleksi fitur yang digunakan adalah *Information Gain*
4. Alat bantu yang digunakan dalam penelitian ini adalah aplikasi *WEKA*.

1.4 TUJUAN DAN MANFAAT PENELITIAN

1.4.1 Tujuan Penelitian

Berdasarkan permasalahan yang tertera di atas, maka dapat disimpulkan bahwa tujuan dari penelitian ini adalah:

1. Untuk menghasilkan data klasifikasi penyakit ginjal kronik menggunakan algoritma *Naïve Bayes* dengan metode seleksi fitur *Information Gain*
2. Dapat mengetahui keakurasian klasifikasi menggunakan algoritma *Naïve Bayes* dengan metode seleksi fitur *Information Gain*
3. Untuk mengoptimasi hasil akurasi dari penelitian sebelumnya dengan menggunakan metode *Naive Bayes* dan teknik seleksi fitur *Information Gain*

1.4.2 Manfaat Penelitian

Manfaat yang akan di dapat dari penelitian ini, yaitu:

1. Dapat digunakan sebagai acuan untuk melakukan penelitian berikutnya dan pengembangan sistem yang jauh lebih baik lagi.
2. Dapat mengoptimalkan hasil akurasi dari data penyakit Ginjal kronik
3. Dapat melakukan pengolahan data dan menghasilkan informasi yang baik

1.5 SISTEMATIKA PENULISAN

Laporan penelitian ini dibuat dalam sistematika yang sesuai dengan kaidah penulisan ilmiah yang benar dan dibagi dalam bab-bab sebagai berikut:

BAB I : PENDAHULUAN

Pada bab ini merupakan bab pendahuluan yang mengemukakan latar belakang masalah, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB II : LANDASAN TEORI

Pada bab ini merupakan landasan teori yang mendasari pembahasan laporan secara khusus berisi definisi dan teori-teori yang dipakai untuk penelitian yang didapat dengan melakukan studi pustaka sebagai dasar dalam melakukan analisis dan perancangan.

BAB III : METODOLOGI PENELITIAN

Pada bab ini berisikan kerangka kerja penelitian, data yang digunakan, alur penelitian, serta alat bantu penelitian yang digunakan untuk mengembangkan proses perhitungan yang dilakukan.

BAB IV : ANALISIS

Pada bab ini menjelaskan tentang bagaimana gambaran umum objek penelitian, analisis sistem yang berjalan, representasi data, pemilihan atribut, dan hasil analisa.

BAB V : KESIMPULAN

Pada bab ini berisi tentang hasil kesimpulan dari perhitungan berdasarkan hasil uji data yang telah dilakukan.