

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Ginjal merupakan salah satu organ sistem saluran kemih atau *urine (traetsu urinalius)* yang berperan menyaring dan mengeluarkan cairan serta sisa metabolisme dari dalam tubuh. Apabila fungsi ginjal menurun secara bertahap mengakibatkan kerusakan ginjal seiring berjalannya waktu yang disebut juga dengan penyakit ginjal kronis. Penyakit Ginjal Kronis (PGK) adalah suatu keadaan yang ditandai dengan meningkatnya prevalensi dan insiden gagal ginjal ditambah dengan prospek yang tidak menjanjikan, dimana biaya yang dikeluarkan untuk pengobatan ini merupakan yang paling besar di seluruh dunia. Penyakit Ginjal Kronis (PGK) atau dalam bahasa Inggris *Chronic Kidney Disease (CKD)* adalah gangguan heterogen yang secara bertahap memengaruhi struktur dan fungsi ginjal, sulit untuk pulih, dan menyebabkan tubuh tidak dapat menjaga metabolisme serta gagal dalam menjaga keseimbangan cairan dan elektrolit, yang berujung pada peningkatan kadar ureum [1].

Berdasarkan hasil laporan dari *World Health Organization (WHO)* tahun 2016, salah satu dari sepuluh penyebab utama kematian di negara-negara berpenghasilan tinggi adalah penyakit ginjal. Tidak hanya itu, penyakit ginjal juga terkait dengan kasus penyakit yang menyebabkan beban kesehatan ekstrem di negara-negara berpendapatan rendah dan menengah, melibatkan sekitar 188 juta

penduduk [2]. Di Indonesia sendiri diperkirakan terdapat 70.000 kasus penderita penyakit ginjal

kronis (PGK) dan merupakan penyakit yang menduduki peringkat kedua dalam pembiayaan terbesar dari BPJS kesehatan setelah penyakit jantung serta untuk mencegah hasil negatif yang terkait dengan PGK termasuk penyakit kardiovaskular, penyakit ginjal stadium akhir ataupun kematian yang memerlukan skrining, diagnosis dan manajemen yang sesuai dibawah bimbingan dokter layanan primer [3]. Banyak individu yang tidak meyakini bahwa mereka memiliki penyakit ginjal kronis mengakibatkan penundaan dalam diagnosa dan akhirnya dapat mengancam nyawa mereka serta rumah sakit membutuhkan manajemen yang baik dalam mengolah data yang sudah ada.

Kemajuan teknologi di bidang kesehatan saat ini mengalami perkembangan yang sangat cepat. Muncul inovasi baru yang mencakup berbagai hal, mulai dari klasifikasi penyakit berdasarkan diagnosis hingga sistem pendukung keputusan yang dapat membantu dalam memberikan rekomendasi yang mendukung prediksi penyakit. Meskipun teknologi ini bukan pengganti keputusan akhir dalam mendiagnosa akan tetapi ia berfungsi sebagai alat bantu yang sangat berharga dalam mendukung proses prediksi penyakit dan pengambilan keputusan dibidang kesehatan. Perkembangan teknologi tersebut dapat menggunakan berbagai algoritma *machine learning*.

Machine learning sendiri merupakan ilmu yang mempelajari tentang algoritma komputer yang bisa mengenali pola-pola di dalam data, dengan tujuan untuk mengubah beragam data menjadi suatu tindakan yang nyata dengan sesedikit mungkin campur tangan manusia [4]. Secara umum *machine learning* terbagi menjadi 2 kategori yaitu *supervised* dan *unsupervised learning*, metode

supervised merupakan metode yang datanya sudah memiliki label, metode ini digunakan untuk membuat prediksi atau klasifikasi pada data baru sedangkan *unsupervised* adalah metode yang tidak terdapat label di dalam datanya. Di dalam *supervised learning* sendiri ada berbagai macam algoritma khususnya *Naïve Bayes Classifier* (NBC) dan *Random Forest* yang merupakan kumpulan dari *Decision tree*.

Naïve Bayes Classifier (NBC) digunakan untuk mengambil keputusan dari sebuah informasi atau data, dimana jumlah data pelatihan (training data) yang tidak terlalu besar sehingga cukup mudah dalam menentukan perkiraan parameter yang dibutuhkan dalam klasifikasi datanya serta cara ini juga banyak digunakan terutama dalam mendiagnosa terkait probabilitas secara statistik dan kemungkinan gejala dan penyakit terkait berdasarkan pengalaman di masa sebelumnya [5]. Sedangkan *Random Forest* sendiri merupakan metode yang dikembangkan dari 2 gabungan jenis pohon yaitu *Classification* dan *Regression Tree* (CART) dengan menerapkan *random feature selection* dan teknik *bagging* yang dimana data training dipilih secara random berdasarkan distribusi vektor acak X, Y [6].

Dari penjelasan diatas belum diketahui algoritma apa yang terbaik dalam mengklasifikasi penyakit ginjal kronis (PGK) diantara kedua algoritma tersebut sehingga dilakukannya perbandingan algoritma dalam klasifikasi penyakit ginjal kronis (PGK) antara *Naïve Bayes Classifier* (NBC) dengan *Random Forest*.

1.2 PERUMUSAN MASALAH

Berdasarkan uraian dari latar belakang diatas maka masalah yang dapat dirumuskan dalam penelitian ini yaitu :

1. Bagaimana menerapkan *Naïve Bayes Classifier* (NBC) dan *Random Forest* sebagai pendukung dalam memprediksi serta pengklasifikasian penyakit ginjal kronis (PGK)?
2. Bagaimana perbandingan performa kedua algoritma *Naïve Bayes Classifier* (NBC) dan *Random Forest* yang lebih akurat dalam mengidentifikasi penyakit ginjal kronis pada pasien?

1.3 BATASAN MASALAH

Agar tidak mengarah pada pembahasan diluar dari penelitian ini, maka perlu dilakukan pembatasan masalah yang akan diteliti, yaitu :

1. Penelitian ini menggunakan data yang didapatkan dari Kaggle dengan alamat website <https://www.kaggle.com/datasets/mansoordaku/ckdisease> yang dimana memiliki 25 atribut diseleksi menjadi 14 atribut dengan total keseluruhan data berjumlah 400 data.
2. Menggunakan algoritma *Naïve Bayes Classifier* dan *Random Forest* untuk menentukan tingkat akurasi serta membandingkan kedua performa algoritma tersebut yang nanti-nya akan menjadi algoritma terbaik dalam mengklasifikasi penyakit ginjal kronis.
3. Alat bantu analisa menggunakan Google Colab.

1.4 TUJUAN DAN MANFAAT PENELITIAN

1.4.1 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dipaparkan, maka tujuan dari penelitian ini adalah :

1. Membuat model algoritma *Naïve Bayes Classifier* (NBC) dan *Random Forest* untuk memperoleh akurasi yang akurat dalam mengklasifikasi penyakit ginjal kronis.
2. Membandingkan algoritma *Naïve Bayes Classifier* (NBC) dengan *Random Forest* untuk mendapatkan performa terbaik diantara kedua algoritma tersebut.

1.4.2 Manfaat Penelitian

Berdasarkan tujuan penelitian diatas dapat diketahui manfaat dari penelitian ini adalah :

1. Sebagai informasi beberapa pihak yang ingin melakukan prediksi maupun klasifikasi penyakit ginjal kronis menggunakan algoritma *Naïve Bayes Classifier* (NBC) dan *Random Forest* untuk mengetahui akurasi yang akurat.
2. Sarana informasi untuk masyarakat yang ingin mengetahui pentingnya kesehatan ginjal dan diharapkan adanya penelitian ini dapat dijadikan bahan perbandingan atau acuan untuk penelitian selanjutnya.

1.5 SISTEMATIKA PENULISAN

Sistematika penulisan tugas akhir penelitian ini disusun untuk memberikan gambaran umum tentang penelitian yang dijalankan. Sistematika penulisan tugas akhir ini adalah sebagai berikut :

BAB I : PENDAHULUAN

Menguraikan tentang latar belakang masalah, merumuskan inti permasalahan yang dihadapi, menentukan maksud dan tujuan dari penelitian yang dilakukan kemudian diikuti dengan pembatasan masalah serta sistematika penulisan.

BAB II : LANDASAN TEORI

Bab ini berfungsi sebagai landasan penyusunan hipotesis penelitian untuk menunjukkan kebenaran teori yang digunakan, sebagai landasan untuk menunjang pemahaman terhadap penelitian yang dilakukan. Landasan teori sendiri bertujuan untuk mengumpulkan semua dokumen yang diperlukan untuk penelitian sebagai landasan yang kuat seperti algoritma, klasifikasi, *naïve bayes classifier*, *random forest*, *machine learning*, dan alat bantu analisis data.

BAB III : METODOLOGI PENELITIAN

Pada bab ini akan menjelaskan tentang kerangka kerja penelitian, pengumpulan data dan metode yang digunakan serta alat bantu (tools) yang digunakan dalam melakukan penelitian ini.

BAB IV : ANALISIS DAN HASIL

Pada bab ini berisi tentang analisis data yang sudah didapatkan lalu dilakukannya perhitungan menggunakan metode *Naïve Bayes Classifier* (NBC) dan *Random Forest* serta hasil dari analisis maupun perhitungan yang akan dijadikan perbandingan diantara kedua algoritma tersebut menggunakan *python*.

BAB V : PENUTUP

Bab ini merupakan bab penutup yang berisikan kesimpulan dari keseluruhan uraian pekerjaan yang telah dilakukan pada bab-bab sebelumnya serta saran yang diperlukan.