

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Penyakit ginjal kronis merupakan masalah kesehatan global yang mempengaruhi banyak orang di seluruh dunia, dengan angka kasus yang terus meningkat[1]. Penyakit ginjal kronis adalah kondisi di mana fungsi ginjal secara bertahap menurun dan tidak dapat pulih. Penyebab utamanya meliputi diabetes, tekanan darah tinggi, serta faktor genetik dan lingkungan. Penyakit ginjal kronis dapat berkembang menjadi gagal ginjal terminal atau stadium akhir, di mana ginjal tidak lagi mampu berfungsi untuk membersihkan tubuh dari zat-zat berbahaya. Pada tahap ini, diperlukan langkah-langkah pemulihan lebih lanjut seperti dialisis atau transplantasi ginjal sebagai pengganti fungsi ginjal.[2]. Pada tahap stadium awal penyakit ini tidak menunjukkan gejala, akan tetapi dapat berkembang dengan perlahan dan menjadi masalah serius tanpa terdeteksi [3].

Salah satu penanganan gagal ginjal kronis adalah hemodialisis (HD), yang harus dilakukan setiap dua kali seminggu dengan setiap sesi berdurasi antara 4 hingga 5 jam. HD bisa menyebabkan efek samping seperti kulit kering dan gatal, kram kaki, pembengkakan, kesulitan tidur, kurang nafsu makan, serta mudah lelah saat beraktivitas. Efek-efek ini dapat mempengaruhi citra tubuh, kualitas hidup, dan kualitas tidur pasien[4]. Setiap masalah dalam hemodialisis dapat menyebabkan ketidaknyamanan serta penurunan kualitas hidup, yang meliputi aspek kesehatan fisik, fisiologis, psikologis, dan status psikososial[1]. Dengan menganalisis data untuk menyesuaikan pengobatan bagi setiap pasien berdasarkan karakteristik individual dan respons mereka terhadap terapi sebelumnya.. [5]. Pendekatan data science ini menawarkan berbagai solusi salah satunya tidak hanya memperbaiki diagnosis dan pengobatan gagal ginjal kronis tetapi juga membantu mengelola sumber daya kesehatan dengan lebih efisien serta meningkatkan kualitas hidup pasien [6].

Dalam penelitian ini, pendekatan yang diterapkan adalah K-Nearest Neighbor (KNN) dan Naive Bayes. Algoritma K-Nearest Neighbor (KNN) adalah salah satu pendekatan untuk klasifikasi yang mengukur jarak antara data uji dan data pelatihan [7]. Metode K-Nearest Neighbor (KNN) mampu mengklasifikasikan data yang sudah ada ke dalam beberapa kelas.

Algoritma K-Nearest Neighbor (KNN) digunakan untuk melakukan klasifikasi objek yang berdasarkan atribut dari data uji dan data latih [8]. Algoritma K-Nearest Neighbor juga sering digunakan untuk klasifikasi dalam pengenalan pola konsistensi data, [9]. Metode *K-Nearest Neighbor* (KNN) adalah metode klasifikasi yang memiliki banyak kelebihan, salah satunya adalah penerapannya yang mudah tetapi tetap efektif dalam berbagai situasi [10].

Algoritma Naïve Bayes, diusulkan oleh ilmuwan Inggris Thomas Bayes, adalah metode klasifikasi dan prediksi yang menggunakan statistika dan probabilitas untuk menganalisis data masa lalu guna memprediksi peluang kejadian di masa depan[11]. Algoritma ini berasumsi bahwa setiap kejadian tidak saling bergantung satu sama lain dan menghitung probabilitas dari setiap variabel untuk membuat prediksi akurat. [12][13].

Berdasarkan pemaparan diatas maka peneliti tertarik melakukan penelitian dengan judul Perbandingan Algoritma *K-Nearest Neighbor* Dan *Naive Bayes* Dalam Prediksi Penyakit Ginjal Kronis Pada Lansia.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, rumusan masalah pada penelitian ini adalah:

1. Bagaimana kinerja dari algoritma *K-Nearest Neighbor* dan *Naive Bayes* dalam melakukan prediksi penyakit ginjal kronis pada lansia?
2. Dari kedua algoritma yang digunakan pada penelitian ini yaitu *K-Nearest Neighbor* dan *Naive Bayes* manakah yang memiliki kinerja terbaik?

1.3 Tujuan dan Manfaat Penelitian

1.3.1 Tujuan Penelitian

Tujuan yang diharapkan dari penelitian ini yaitu:

1. Untuk mengetahui kinerja dari algoritma *K-Nearest Neighbor* dan *Naive Bayes* dalam melakukan prediksi penyakit ginjal kronis pada lansia.
2. Untuk mengetahui dari ketiga algoritma data *mining* yang dibahas pada penelitian ini yaitu *K-Nearest Neighbor* dan *Naive Bayes* manakah yang memiliki kinerja terbaik.

1.3.2 Manfaat Penelitian

Manfaat yang diharapkan pada penelitian ini yaitu:

1. Pada bidang kesehatan penelitian ini bertujuan untuk menciptakan sistem prediksi yang lebih akurat berdasarkan hasil analisis yang ditemukan.
2. Hasil penelitian ini dapat dijadikan referensi bagi peneliti selanjutnya yang membahas prediksi penyakit dengan menggunakan algoritma *K-Nearest Neighbor* dan *Naive Bayes*.

1.4 Batasan Masalah

Berikut ini akan diuraikan batasan masalah pada penelitian ini agar penelitian berjalan sesuai dengan perencanaan yaitu:

1. Dataset pada penelitian ini diambil dari Kaggle *Chronic Kidney Disease Database* sebanyak 400 data, dari dataset ini peneliti mengambil data berdasarkan usia yaitu 46-65 tahun sehingga diperoleh data sebanyak 195 data.
2. Atribut-atribut *dataset* yang digunakan pada penelitian ini terdiri dari *blood pressure (bp)*, *specific gravity (sg)*, *albumin (al)*, *sugar (su)*, *red blood cell (rbc)*, *pus cell (pc)*, *puss cell clumps (pcc)*, *bacteria (ba)*, *blood glucose random (bgr)*, *blood urea (bu)*, *serum creatinine (sc)*, *sodium (sod)*, *potassium (pot)*, *hemoglobin (hemo)*, *packed cell volume (pcv)*, *white blood cell count (wbcc)*, *red blood cell count (rbcc)*, *hypertension (htn)*, *diabetes mellitus (dm)*.
3. *Output* hasil prediksi terbagi menjadi 2 yaitu *ckd* (menderita penyakit ginjal kronis) dan *notckd* (tidak menderita penyakit ginjal kronis).

1.5 Keterbaruan Penelitian

Berikut ini akan diuraikan keterbaruan dari penelitian ini melalui literature review penelitian terdahulu yang berkaitan dengan topik penelitian antara lain:

1. Menurut Toni Arifin dan Daniel Ariesta (2019) dalam penelitiannya yang berjudul *Prediksi Penyakit Ginjal Kronis Menggunakan Algoritma Naive Bayes Classifier*

Berbasis Particle Swarm Optimization. Berdasarkan penelitian, metode Naive Bayes Classification yang didukung oleh Particle Swarm Optimization menunjukkan akurasi confusion matrix sebesar 98,75% dan AUC mencapai 99%. Sementara itu, metode Naive Bayes tanpa optimisasi menunjukkan akurasi confusion matrix sebesar 97,00% dan AUC sebesar 99,8%.

2. Menurut Shofi Inayah Putri, Tri Kesuma Dewi, Ludiana (2023) dengan judul penelitian Penerapan Slow Deep Breathing Terhadap Kelelahan (Fatigue) Pada Pasien Gagal Ginjal Kronik Di Ruang Hd Rsud Jendral Ahmad Yani Metro Tahun 2022. Hasil penerapan menunjukkan bahwa setelah melakukan slow deep breathing, skor kelelahan pada kedua subjek meningkat dari 32 menjadi 40 dan dari 33 menjadi 39. Bagi pasien dengan gagal ginjal kronis yang menjalani hemodialisis, disarankan untuk secara rutin dan mandiri melakukan slow deep breathing, karena dapat membantu mengurangi kelelahan..
3. Siti Noor Chotimah, Ade Ricky Rozzaqi (2023) dalam penelitiannya yang berjudul Klasifikasi Diagnosis Penyakit Ginjal Kronis Dengan Menerapkan Konsep Algoritma Naïve Bayes. Penelitian ini bertujuan untuk mengklasifikasikan penyakit ginjal kronis menggunakan algoritma Naïve Bayes dan untuk mengidentifikasi faktor-faktor yang mempengaruhi tingkat akurasi. Pada tahap pertama, dataset mentah diproses melalui pre-processing untuk menangani missing value dan melakukan transformasi data. Pada tahap kedua, dataset yang telah diproses diklasifikasikan dengan dua pendekatan. Pendekatan pertama menggunakan cross validation dengan $k=5$, sementara pendekatan kedua menggunakan $k=10$. Hasil menunjukkan bahwa pendekatan dengan $k=5$ menghasilkan akurasi yang lebih tinggi, yaitu 80%, dibandingkan dengan $k=10$ yang menghasilkan akurasi sebesar 76%. Penggunaan $k=5$ dianggap lebih efektif dalam mengurangi kemungkinan overfitting pada dataset yang digunakan. Tingkat akurasi yang diperoleh dipengaruhi oleh perlakuan yang diterapkan pada dataset selama penelitian.
4. Resa Rostira Pebriani (2024) dalam penelitiannya yang berjudul Diagnosis Penyakit Ginjal Kronis (PGK) Menggunakan Metode Klasifikasi Naive Bayes. Dalam penelitian ini, data rekam medis elektronik yang digunakan berasal dari dataset UCI Machine Learning. Dataset tersebut terdiri dari 400 data dan 24 fitur. Hasil evaluasi model klasifikasi menunjukkan akurasi 97%, presisi 97%, recall 97%, dan F1-score

97%. Berdasarkan nilai-nilai tersebut, model klasifikasi yang dikembangkan menunjukkan performa yang sangat baik dalam mengklasifikasikan data. Diagnosa Penyakit Ginjal Kronis (PGK) dapat dilakukan menggunakan antarmuka sederhana yang berbasis aplikasi web.