



# Perbandingan Algoritma K-Nearest Neighbor dan Naive Bayes Dalam Prediksi Penyakit Ginjal Kronis Pada Lansia

Marco Duran Simbolon\*, Dimas Dimanta Bukit, Rian Elby Purba, Faisal Haries Ketaren,  
Agung Prabowo

Fakultas Sains dan Teknologi, Program Studi Teknik Informatika, Universitas Prima Indonesia, Medan

Jl. Sampul No.3, Sei Putih Barat, Kec. Medan Petisah, Kota Medan, Sumatera Utara, Indonesia

Email: <sup>1,\*</sup>marcooduran2001@gmail.com, <sup>2</sup>dimasbukit25@gmail.com, <sup>3</sup>rianpurba2000@gmail.com,

<sup>4</sup>faisalketaren77@gmail.com, <sup>5</sup>agungprabowo@unprimdn.ac.id

Email Penulis Korespondensi: marcooduran2001@gmail.com

Submitted: 20/09/2024; Accepted: 31/01/2025; Published: 31/01/2025

**Abstrak**—Penyakit ginjal kronis merupakan salah satu penyakit serius yang memerlukan diagnosis dini untuk meningkatkan peluang perawatan yang lebih baik, terutama pada lansia. Permasalahan utama dalam diagnosis penyakit ini adalah sering kali gejala tidak tampak hingga penyakit berada pada tahap lanjut, sehingga menuntut adanya metode prediksi yang akurat. Selain itu, keterbatasan ukuran dataset yang hanya terdiri dari 195 data pasien dapat mempengaruhi kemampuan algoritma dalam mengenali pola yang ada. Pemilihan algoritma yang tepat juga menjadi tantangan karena beberapa algoritma memiliki keterbatasan dalam menangani data medis yang kompleks. Penelitian ini bertujuan untuk mengevaluasi kemampuan algoritma K-Nearest Neighbors (KNN) dan Naive Bayes dalam memprediksi penyakit ginjal kronis. Dataset dianalisis menggunakan perangkat lunak Weka Waikato dan diuji dengan metode k-fold cross-validation sebesar 9. Hasil terbaik diperoleh pada pengujian menggunakan algoritma Naive Bayes, dengan akurasi sebesar 97,4359%. Berdasarkan hasil tersebut, dapat disimpulkan bahwa kedua algoritma dapat digunakan untuk prediksi penyakit ginjal kronis pada lansia. Namun, solusi yang dapat diusulkan untuk meningkatkan akurasi prediksi lebih lanjut meliputi penambahan dataset yang lebih besar dan beragam serta pengoptimalan hyperparameter algoritma. Di sisi lain, algoritma Naive Bayes menunjukkan tingkat akurasi yang lebih tinggi dibandingkan KNN dalam penelitian ini, sehingga lebih direkomendasikan untuk digunakan..

**Kata Kunci:** Penyakit Ginjal Kronis; K-Nearest Neighbors; Naive Bayes; Prediksi; Lansia

**Abstract**—Chronic kidney disease is a serious illness that requires early diagnosis to improve treatment outcomes, especially in the elderly. The main challenge in diagnosing this disease lies in the fact that symptoms often do not appear until the disease has reached an advanced stage, which necessitates the use of accurate prediction methods. Additionally, the dataset's limited size, consisting of only 195 patient records, may affect the algorithm's ability to identify patterns. Choosing the appropriate algorithm is also a challenge, as some algorithms have limitations in handling complex medical data. This study aims to evaluate the performance of the K-Nearest Neighbors (KNN) and Naive Bayes algorithms in predicting chronic kidney disease. The dataset was analyzed using Weka Waikato software and tested using the 9-fold cross-validation method. The best results were obtained using the Naive Bayes algorithm, with an accuracy of 97.4359%. Based on these results, it can be concluded that both algorithms can be used to predict chronic kidney disease in the elderly. However, to further improve prediction accuracy, proposed solutions include expanding the dataset with more diverse data and optimizing the algorithm's hyperparameters. On the other hand, the Naive Bayes algorithm demonstrated higher accuracy compared to KNN in this study, making it the more recommended choice.

**Keywords:** Chronic Kidney Disease; K-Nearest Neighbors; Naive Bayes; Prediction; Elderly

## 1. PENDAHULUAN

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].

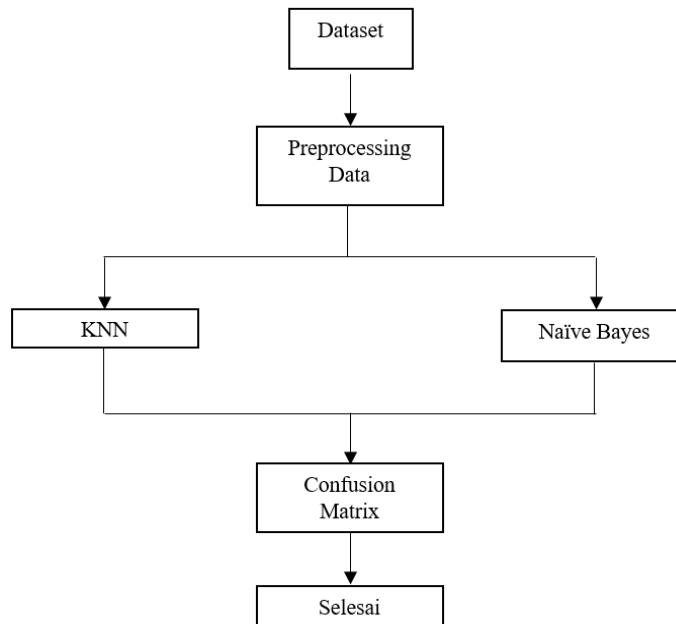
Dalam penelitian ini, pendekatan yang diterapkan adalah K-Nearest Neighbor (KNN) dan Naive Bayes. Algoritma K-Nearest Neighbor (KNN) adalah salah satu pendekatan untuk klasifikasi yang mengukur jarak antara

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.

## 2. METODOLOGI PENELITIAN

Pada bagian ini peneliti peneliti menggambarkan langkah-langkah dalam tahapan penelitian yang dilakukan. Tahapan penelitian ini dapat dilihat pada Gambar 1 berikut.



**Gambar 1.** Tahapan Penelitian

### 2.1 Pengumpulan dataset

Pada penelitian ini dataset diperoleh dari kaggle yaitu dataset penyakit ginjal kronis dengan jumlah data sebanyak 400 data dengan atribut kondisi 24 dan atribut class 1. Pada dataset ini rentang usia masih umum sehingga dilakukan pembersihan data dengan membuang data yang tidak masuk pada usia lansia yaitu rentang usi dari 45 – 65 tahun. Dari rentang usia lansia diperoleh data berjumlah 195 data. Rincian dataset Chronic Kidney Disease Lansia pada penelitian ini dapat dilihat pada Tabel 1 berikut.

**Tabel 1.** Dataset Penyakit Ginjal Kronis Lansia

| no  | id  | age | bp  | sg   | ... | appet | pe  | ane | classification |
|-----|-----|-----|-----|------|-----|-------|-----|-----|----------------|
| 1   | 41  | 45  | 70  | 1,01 | ... | no    | yes | no  | ckd            |
| 2   | 67  | 45  | 80  | 1,02 | ... |       |     | no  | ckd            |
| 3   | 81  | 45  | 90  | 1,01 | ... |       | no  |     | ckd            |
| 4   | 108 | 45  |     |      | ... |       |     |     | ckd            |
| 5   | 137 | 45  |     |      | ... | yes   |     | no  | ckd            |
| 6   | 181 | 45  | 70  |      | ... | yes   | no  | no  | ckd            |
| ... | ... | ... | ... | ...  | ... | ...   | ... | ... | ...            |
| 193 | 71  | 65  | 60  | 1,25 | ... | yes   | no  | yes | notckd         |
| 194 | 117 | 65  |     |      | ... |       | no  | no  | ckd            |
| 195 | 132 | 65  |     |      | ... |       |     |     | notckd         |
| 196 | 181 | 65  | 80  | 1,12 | ... | yes   | no  | yes | ckd            |



### 2.2 Preprocessing Data

Pada tahap ini peneliti melakukan normalisasi data agar fitur-fitur memiliki skala nilai yang sama. Pada dataset yang ada proses normalisasi menggunakan metode one hot encoding yaitu mengubah setiap nilai kategori ke dalam bentuk biner atau 0 dan 1. Fitur yang dinormalisasi adalah rbc, pc, pcc, ba, htn, dm, cad, appet, pe, ane, dan class. Tabel 2 berikut adalah encoding untuk setiap fitur.

Tabel 2. Hasil Encoding Dataset Penyakit Ginjal Kronis Lansia

| no  | id  | age | Bp  | sg    | al  | Su  | ... | ... | ... | dm  | cad | appet | pe  | ane | classification |
|-----|-----|-----|-----|-------|-----|-----|-----|-----|-----|-----|-----|-------|-----|-----|----------------|
| 1   | 41  | 45  | 70  | 1,01  | 0   | 0   | ... | ... | ... | 0   | 0   | 0     | 1   | 0   | 1              |
| 2   | 67  | 45  | 80  | 1,02  | 3   | 0   | ... | ... | ... | 1   | 0   | 0     | 0   | 0   | 1              |
| 3   | 81  | 45  | 90  | 1,01  | 3   | 1   | ... | ... | ... | 1   | 0   | 1     | 0   | 0   | 1              |
| 4   | 108 | 45  | 80  | 1,015 | 0   | 0   | ... | ... | ... | 0   | 0   | 0     | 0   | 0   | 1              |
| 5   | 137 | 45  | 60  | 1,01  | 2   | 0   | ... | ... | ... | 1   | 0   | 1     | 0   | 0   | 1              |
| 6   | 181 | 45  | 70  | 1,025 | 2   | 0   | ... | ... | ... | 1   | 0   | 1     | 0   | 0   | 1              |
| ... | ... | ... | ... | ...   | ... | ... | ... | ... | ... | ... | ... | ...   | ... | ... | ...            |
| 193 | 71  | 45  | 60  | 1,25  | 2   | 0   | ... | ... | ... | 0   | 0   | 1     | 0   | 1   | 0              |
| 194 | 117 | 45  | 60  | 1,01  | 3   | 1   | ... | ... | ... | 1   | 0   | 0     | 0   | 0   | 1              |
| 195 | 132 | 45  | 70  | 1,025 | 0   | 1   | ... | ... | ... | 1   | 0   | 1     | 0   | 1   | 0              |
| 196 | 181 | 45  | 80  | 1,12  | 2   | 1   | ... | ... | ... | 0   | 0   | 1     | 0   | 1   | 1              |

### 2.4 Algoritma KNN

K-Nearest Neighbour (K-NN) adalah metode yang digunakan untuk mengklasifikasikan objek baru berdasarkan (k) tetangga terdekatnya [14]. K-NN adalah algoritma supervised learning, di mana hasil klasifikasi instance baru ditentukan berdasarkan mayoritas kategori dari K tetangga terdekatnya. Kelas yang paling sering muncul di antara tetangga tersebut akan menjadi kelas hasil klasifikasi [15]. Masalah yang dapat diselesaikan oleh K-NN adalah ketika terdapat data training dalam jumlah besar, algoritma ini tangguh dan efektif dalam menangani data yang banyak. Langkah-langkah dalam menentukan algoritma K-NN meliputi: menentukan parameter, menghitung jarak antara data yang dievaluasi dengan seluruh data latih, mengurutkan jarak terdekat, menentukan tetangga terdekat, memasang kelas yang sesuai, dan akhirnya, menghitung jumlah kelas terbanyak dari tetangga terdekat tersebut dan menetapkannya sebagai kelas data yang sedang dievaluasi [16].

### 2.5 Algoritma Naive Bayes

Metode Naive Bayes adalah teknik klasifikasi atau pengelompokan data yang menggunakan perhitungan probabilitas berdasarkan data yang telah terkumpul [17]. Naive Bayes digunakan untuk memprediksi peluang di masa depan dengan memanfaatkan metode probabilitas serta analisis statistik berdasarkan data dari pengalaman sebelumnya [18]. Salah satu kelebihan utama dari algoritma Naive Bayes adalah kebutuhan jumlah data pelatihan yang relatif sedikit untuk mengestimasi parameter dalam proses klasifikasi. Selain itu, penerapan Naive Bayes cukup sederhana dan sering menghasilkan performa yang baik di berbagai kondisi [19]. Namun, kekurangan dari metode ini adalah asumsi independensi antar fitur. Dalam kenyataannya, keterkaitan antar fitur seringkali ada dan tidak dapat dimodelkan oleh NaiveBayesian Classifier. Persamaan dari algoritma naive bayes dapat dilihat pada persamaan berikut [20].

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)} \tag{1}$$

C<sub>i</sub> merupakan hipotesis data X yang merupakan kelas yang spesifik. Pada klasifikasi naive bayes C<sub>i</sub> adalah salah satu kelas dari sekumpulan teks yang ada, seperti spam atau no spam dalam klasifikasi email. X merupakan data dengan kelas yang belum diketahui, contohnya adalah pada klasifikasi X adalah atribut atau fitur dari data, seperti kata-kata dalam email yang ingin diklasifikasikan. P(C<sub>i</sub>|X) merupakan probabilitas hipotesis C<sub>i</sub> berdasarkan kondisi X. Ini juga disebut dengan probabilitas posterior yaitu probabilitas bahwa hipotesis C<sub>i</sub> benar, diberikan data X. P(X|C<sub>i</sub>) merupakan probabilitas X berdasarkan kondisi pada hipotesis C<sub>i</sub>. Ini juga disebut probabilitas likelihood yaitu probabilitas munculnya data X jika diketahui bahwa hipotesis C<sub>i</sub> benar. P(C<sub>i</sub>) atau probabilitas prior yaitu probabilitas awal dari hipotesis C<sub>i</sub> sebelum data X diamati. Contohnya adalah probabilitas awal bahwa sebuah email adalah spam berdasarkan pengalaman sebelumnya. P(X) atau probabilitas evidence yaitu probabilitas data X terlepas dari kelas C<sub>i</sub>. Probabilitas evidence ini merupakan faktor normalisasi agar probabilitas posterior menjadi nilai yang valid.

### 2.6 Confusion Matrix

Tahapan selanjutnya yaitu merupakan proses pembagian data latih dan ujiselesai, evaluasi model dilakukan. Penelitian ini dalam melakukan evaluasi keakuratan model, evaluasi dilakukan dengan menggunakan confusion matrix, digunakan rumus confusion matrix. confusion matrix memberikan wawasan yang lebih mendalam

mengenai performa model dalam mengklasifikasikan kelas minoritas maupun mayoritas [21] [22]. Persamaan dari confusion matrix dapat dilihat pada persamaan berikut ini [23].

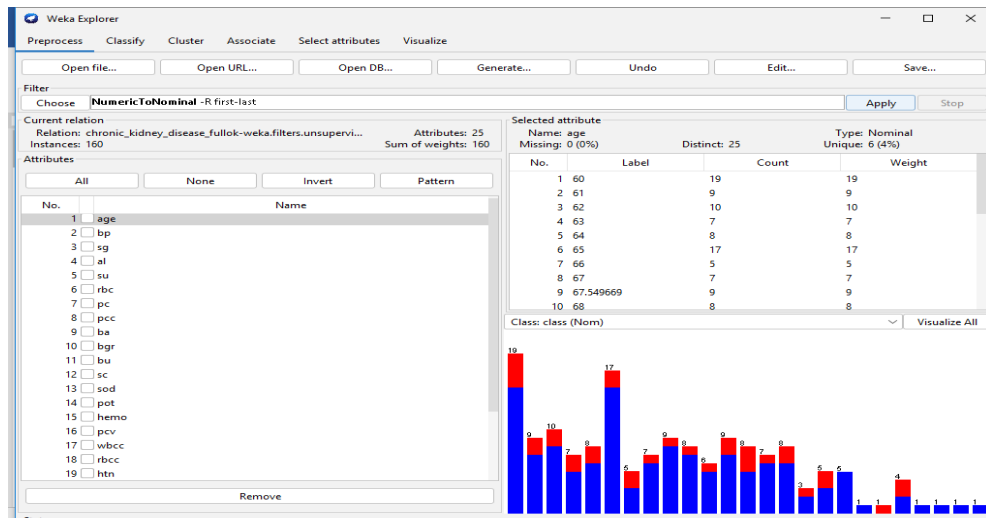
$$\text{Accuracy} = \frac{(TP+TN)}{(TN+TP+FN+FP)} \tag{2}$$

True Positive (TP): merupakan jumlah tebakan yang benar, di mana model kita menebak seseorang terinfeksi penyakit ginjal kronis, dan kenyataannya memang benar. Misalnya, model menebak seseorang positif penyakit ginjal kronis, dan hasil tes memang menunjukkan positif penyakit ginjal kronis. True Negative (TN): merupakan jumlah tebakan yang benar, di mana model kita menebak seseorang tidak terinfeksi penyakit ginjal kronis, dan kenyataannya memang benar. Misalnya, model menebak seseorang negatif penyakit ginjal kronis, dan hasil tes memang menunjukkan negatif penyakit ginjal kronis. False Positive (FP): merupakan jumlah tebakan di mana model kita menebak seseorang terinfeksi penyakit ginjal kronis, tapi kenyataannya itu salah. Misalnya, model menebak seseorang positif penyakit ginjal kronis, padahal hasil tes sebenarnya menunjukkan negatif penyakit ginjal kronis. False Negative (FN): merupakan jumlah tebakan di mana model kita menebak seseorang tidak terinfeksi penyakit ginjal kronis, tapi kenyataannya itu salah. Misalnya, model menebak seseorang negatif penyakit ginjal kronis, padahal hasil tes sebenarnya menunjukkan positif penyakit ginjal kronis

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Hasil

Sebelum dilakukan proses algoritma KNN dan Naïve Bayes dalam melakukan prediksi penyakit ginjal kronis pada lansia menggunakan aplikasi weka. Data mentah yang digunakan pada penelitian ini memerlukan beberapa penyesuaian agar dapat diolah secara efektif menggunakan lagortima yang telah dipilih. Salah satu langkah utama yang dilakukan adalah mengubah dataset dari format numerik menjadi nominal dengan tujuan memastikan bahwa setiap atribut pada dataset dapat diinterpretasikan dengan baik oleh algoritma yang membutuhkan data dalam bentuk kategori. Selain itu proses ini dilakukan untuk memastikan konsistensi pada pengolahan data. Dataset mentah biasanya memiliki nilai yang tidak seragam ataupun tidak lengkap sehingga diperlukan pembersihan data sebelum mengubahnya menjadi nominal. Hal ini dapat kita lihat pada Gambar 2 berikut.



**Gambar 2** Perubahan Dataset dari Numerik menjadi Nominal

Setelah dataset seluruhnya menjadi nominal, maka langkah selanjutnya adalah melakukan prediksi menggunakan algoritma KNN dengan cross validation dan nilai k=2, 3, 4, 5, 6, 7, 8, 9 dan 10. Adapun hasil dari penggunaan algoritma KNN dapat dilihat pada Tabel 2 berikut.

**Tabel 2.** Hasil Pengujian Algoritma KNN

| K-Fold | Accuracy | Absolute Errorr | RMSE   |
|--------|----------|-----------------|--------|
| 2      | 94.8718  | 0.078           | 0.2278 |
| 3      | 95.8974  | 0.0597          | 0.193  |
| 4      | 94.359   | 0.0706          | 0.2209 |
| 5      | 94.359   | 0.0622          | 0.2108 |
| 6      | 95.3846  | 0.0604          | 0.2035 |
| 7      | 94.8718  | 0.0639          | 0.215  |
| 8      | 94.359   | 0.0669          | 0.2165 |



| K-Fold | Accuracy | Absolute Errorr | RMSE   |
|--------|----------|-----------------|--------|
| 9      | 95.3846  | 0.0579          | 0.2019 |
| 10     | 94.8718  | 0.0657          | 0.2167 |

Langkah selanjutnya adalah melakukan prediksi menggunakan algoritma Naïve Bayes dengan cross validation dan nilai k=2, 3, 4, 5, 6, 7, 8, 9 dan 10. Adapun hasil dari penggunaan algoritma Naïve Bayes dapat dilihat pada gambar berikut.

**Tabel 3.** Hasil Pengujian Algoritma Naïve Bayes

| K-Fold | Accuracy | Absolute Errorr | RMSE   |
|--------|----------|-----------------|--------|
| 2      | 96.4103  | 0.0419          | 0.1769 |
| 3      | 96.4103  | 0.0378          | 0.1676 |
| 4      | 96.9231  | 0.0356          | 0.169  |
| 5      | 95.8974  | 0.0374          | 0.1669 |
| 6      | 96.9231  | 0.0341          | 0.1593 |
| 7      | 96.9231  | 0.0325          | 0.159  |
| 8      | 96.9231  | 0.0351          | 0.1647 |
| 9      | 97.4359  | 0.0326          | 0.1548 |
| 10     | 96.9231  | 0.0339          | 0.159  |

### 3.2 Pembahasan

Pada algoritma KNN tingkat akurasi terbaik diperoleh pada k-fold sebesar 3 dengan tingkat akurasi yaitu sebesar 95,8974%. Kemudian absolute errorr yaitu 0,0597 dan Root Mean Square (RMSE) yang diperoleh 0,193. Hal ini menunjukkan bahwa KNN memiliki kemampuan yang cukup baik untuk prediksi penyakit ginjal kronis pada lansia, terutama dengan parameter k-fold yang dioptimalkan. Namun nilai k-fold yang berbeda memperoleh variasi performa akurasi, sehingga sangat penting untuk memilih nilai k yang sesuai dengan dataset yang digunakan.

Algoritma KNN bekerja dengan cara melakukan klasifikasi data baru berdasarkan kedekatan dengan data yang sudah ada. KNN sangat bergantung pada nilai k, yaitu jumlah tetangga terdekat yang dipertimbangkan dalam proses klasifikasi. Pada penelitian ini, k-fold yang menunjukkan hasil terbaik adalah pada k-fold sebesar 3 karena memungkinkan algoritma KNN menemukan keseimbangan antara underfitting dan overfitting. Meskipun demikian, algoritma ini memiliki kelemahan yaitu waktu komputasi yang lebih lama untuk menangani dataset yang besar, karena harus menghitung jarak setiap titik data.

Pada algoritma naïve bayes performa ditunjukkan lebih konsisten dan akurat. Performa pada algoritma ini diperoleh pada k-fold sebesar 9, yaitu sebesar 97,4359%, dengan nilai absolute errorr yang dihasilkan lebih rendah dibandingkan algoritma KNN yaitu sebesar 0,0326 dan Root Mean Square Error (RMSE) sebesar 0,1548. Hasil ini menunjukkan bahwa algoritma naïve bayes memiliki kemampuan yang lebih unggul dalam menemukan pola dan hubungan antara atribut pada dataset yang digunakan, meskipun algoritma ini menggunakan pendekatan probabilitas sederhana.

Algoritma naïve bayes mengasumsikan setiap atribut pada dataset bersifat independen. Namun asumsi ini sering tidak realistis dalam dunia nyata karena algoritma ini memberikan hasil yang baik karena kesederhanaannya dalam menangani data. Pada kasus ini, naïve bayes mampu memanfaatkan probabilitas dari dataset penyakit ginjal kronis untuk mendapatkan hasil prediksi yang lebih akurat dibandingkan dengan algoritma KNN.

Kedua algoritma ini memiliki keunggulannya masing-masing, dimana algoritma KNN lebih efektif pada dataset dengan jumlah data yang tidak terlalu besar dan parameter k disesuaikan dengan baik. Sedangkan algoritma naïve bayes memiliki keunggulan dalam menangani dataset dengan jumlah fitur yang lebih banyak karena memiliki efisiensi dalam memproses data. Berdasarkan hal tersebut untuk dataset penyakit ginjal kronis pada lansia algoritma yang lebih baik adalah naïve bayes.

Selain tingkat akurasi, efisiensi waktu komputasi menjadi faktor yang tidak kalah penting dalam memilih algoritma yang digunakan. Dalam hal ini algoritma naïve bayes cenderung lebih cepat dibandingkan dengan algoritma KNN karena proses klasifikasinya tidak memerlukan perhitungan jarak setiap data. Pada konteks klinis, efisiensi menjadi sangat penting karena keputusan cepat seringkali diperlukan untuk diagnosis dini dan pengobatan penyakit ginjal kronis.

Dengan mempertimbangkan seluruh hasil analisis, penelitian ini menyimpulkan bahwa algoritma naïve bayes adalah algoritma yang cocok untuk prediksi penyakit ginjal kronis pada lansia. Algoritma ini unggul tidak hanya pada akurasi namun dalam efisiensi proses sehingga menjadikan algoritma ini pilihan utama untuk studi lanjutan dan aplikasi praktis pada pengelolaan data kesehatan.

## 4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan menggunakan Weka Waikato terhadap 195 data penyakit ginjal kronis, pengujian dilakukan dengan algoritma K-Nearest Neighbors (KNN) dan Naïve Bayes. Hasil terbaik



diperoleh dengan algoritma Naïve Bayes menggunakan 9-fold cross-validation, mencapai akurasi 97,4359%. Kedua algoritma tersebut mampu memprediksi penyakit ginjal kronis pada lansia, namun Naïve Bayes menunjukkan tingkat akurasi yang lebih tinggi dibandingkan KNN. Hasil ini menunjukkan bahwa Naïve Bayes dapat menjadi alat yang lebih andal untuk memprediksi penyakit ginjal kronis, terutama dengan dataset ini. Meskipun KNN dikenal karena kesederhanaan dan efektivitasnya dalam tugas pengenalan pola, dalam kasus ini, Naïve Bayes mengungguli KNN. Pendekatan probabilistik Naïve Bayes terbukti lebih mampu menangkap hubungan mendasar dalam dataset, sehingga menghasilkan prediksi yang lebih akurat. Selain itu, algoritma Naïve Bayes lebih efisien dalam hal waktu komputasi, yang sangat penting dalam aplikasi klinis di mana keputusan cepat sering kali diperlukan. Meskipun KNN dapat memberikan hasil yang baik dalam situasi tertentu, kecenderungan Naïve Bayes untuk memberikan akurasi tinggi membuatnya lebih diunggulkan dalam konteks diagnosis penyakit. Oleh karena itu, untuk penelitian dan praktik di masa depan dalam memprediksi penyakit ginjal kronis, terutama di kalangan lansia, algoritma Naïve Bayes mungkin lebih disukai daripada KNN, khususnya ketika akurasi tinggi menjadi prioritas dalam pengambilan keputusan untuk diagnosis dini dan pengobatan.

## REFERENCES

- [1] T. Pipit Muliyah, Dyah Aminatun, Sukma Septian Nasution, Tommy Hastomo, Setiana Sri Wahyuni Sitepu, "Faktor Yang Berhubungan Dengan Tingkat Kecemasan Pasien Penyakit Ginjal Kronik Yang Menjalani Hemodialisis," *J. GEEJ*, vol. 7, no. 2, pp. 831–844, 2020.
- [2] and Y. R. M. Rizal, M. Z. Syahaf, S. R. Priyambodo, "Optimasi Algoritma Naïve Bayes Menggunakan Forward Selection Untuk Klasifikasi Penyakit Ginjal Kronis," *Naratif J. Nas. Riset, Apl. dan Tek. Inf.*, vol. 5, no. 1, pp. 71–80, 2023.
- [3] P. M. S. Madani, T. Rohana, K. A. Baihaqi, and A. Fauzi, "Perbandingan Kinerja Klasifikasi Penyakit Ginjal Menggunakan Algoritma Support Vector Machine (SVM) dan Decision Tree (DT)," *Build. Informatics, Technol. Sci.*, vol. 6, no. 1, pp. 74–82-74–82, 2024, [Online]. Available: <https://ejurnal.seminar-id.com/index.php/bits/article/view/5206>
- [4] Sutanto, M. Suandika, and A. Susanto, "Hubungan Body Image Dengan Quality of Life Dan Quality of Sleep Pada Pasien Gagal Ginjal Kronik Yang Menjalani Hemodialisa," *J. Penelit. Perawat Prof.*, vol. 6, pp. 976–980, 2024
- [5] and H. P. N. Komariyah, D. N. Aini, "Hubungan Usia, Jenis Kelamin Dan Tingkat Pendidikan Dengan Kepatuhan Pembatasan Cairan Pada Pasien Gagal Ginjal Kronik Yang Menjalani Hemodialisis," *J. Ilm. Permas J. Ilm. STIKES Kendal*, vol. 14, no. 3, pp. 1107–1116, 2024.
- [6] O. V. Hafifi et al., "Analisis Survival Penderita Gagal Ginjal Dengan Pendekatan Multivariate Adaptive Regression Spline," *J. Endur. Kaji. Ilm. Probl. Kesehat.*, vol. 6, no. 3, pp. 504–511, 2021.
- [7] and A. F. V. Alvian, D. Hidayatullah, A. Nilogiri, H. Azizah, "Klasifikasi Siswa Berprestasi Menggunakan Metode K-Nearest Neighbor (KNN) Pada SMA Negeri 2 Situbondo Classification Of Achieving Students Using K-Nearest Neighbor (KNN) Method At SMA Negeri 2 Situbondo," *J. Smart Teknol*, vol. 1, no. 1, pp. 2774–1702, 2021.
- [8] and R. I. N. Fatimah Indrianti, A. Kania Ningsih, "Implementasi Data Mining Untuk Klasifikasi Penyakit Gagal Ginjal Kronis Menggunakan Metode K-Nearest Neighbor," *JATI (Jurnal Mhs. Tek. Inf.*, vol. 8, no. 2, pp. 2255–2260, 2024.
- [9] and P. S. N. N. P. Nugraha, R. Azim, S. Z. Daffa, "Perbandingan Akurasi Metode Naïve Bayes dan Metode KNN untuk Memprediksi Gagal Ginjal Kronis," *J. Rekayasa Elektro Sriwij.*, vol. 5, no. 1, pp. 1–10, 2023.
- [10] I. Wisnuadji Gamadarenda and I. Waspada, "Implementation of Data Mining for the Detection of Chronic Kidney Disease (Ckd) Using K-Nearest Neighbor (Knn) With Backward Elimination," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 2, pp. 417–426, 2020.
- [11] D. Wilandini and P. Purwanto, "Penerapan Algoritma Naïve Bayes dalam mengklasifikasikan Media Sosial untuk mengamati Trend Kuliner," *J. Teknol. Terpadu*, vol. 8, no. 1, pp. 31–39, 2022.
- [12] and M. D. S. N. Rahmadani, R. Risnawati, "Penerapan Algoritma Naïve Bayes Dalam Penentuan Kelayakan Penerima Bantuan Program Keluarga Harapan," *J. Tek.*, vol. 3, no. 2, p. 40, 2023.
- [13] and T. S. I. Hadianti, R. Astuti, "Penerapan Algoritma Naïve Bayes Untuk Menentukan Kelayakan Penerima Bantuan Sosial Di Desa Golat," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 6, pp. 3616–3620, 2024.
- [14] A. Hidayati, S., & Yulianto, "Pengaruh Kesehatan Mental Terhadap Kualitas Hidup Lansia Usia 45-65 Tahun," *J. Kesehat. Masy.*, vol. 16, no. 2, pp. 115–122, 2021.
- [15] N. Prasetyo, E., & Sari, "Tingkat Kecemasan dan Depresi pada Lansia Usia 45-65 Tahun di Era Pandemi COVID-19," *J. Psikol.*, vol. 10, no. 1, pp. 45–54, 2021.
- [16] dan R. H. U. S. A. B. Mukhlis, A. R. Mahendra, "Penanganan Missing Value pada Data Kesehatan Menggunakan Imputasi Rata-rata," *J. Sist. Inf.*, vol. 16, no. 1, pp. 33–40, 2020.
- [17] Alfa Saleh, "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga," *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 1, no. 2, pp. 73–81, 2019.
- [18] S. D. Herry Derajad Wijaya, "Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat," *J. Inform.*, vol. 7, no. 1, pp. 1–7, 2020.
- [19] E. Chely Aulia Misrun, Haerani, E., Fikry, M., & Budianita, "Analisis sentimen komentar youtube terhadap Anies Baswedan sebagai bakal calon presiden 2024 menggunakan metode naive bayes classifier," *J. CoSciTech(Computer Sci. Inf. Technol.*, vol. 4, no. 1, pp. 207–215, 2023.
- [20] W. Johnson, P., & Zhang, "A Comprehensive Analysis of Naïve Bayes Classifier and Its Applications in Machine Learning," *J. Data Sci. Anal.*, vol. 45, no. 3, pp. 207–215, 2023.
- [21] K. Smith, J., & Johnson, "Evaluation Metrics for Machine Learning Models: The Role of Confusion Matrix," *J. Artif. Intell. Res.*, vol. 35, no. 2, pp. 112–130, 2023.
- [22] N. Hidayat, A., & Suryana, "Analisis Kinerja Algoritma Klasifikasi Menggunakan Confusion Matrix pada Dataset



- Kesehatan,” J. Inform. dan Sist. Inf., vol. 15, no. 1, pp. 67–74, 2022.
- [23] F. Prasetyo, A., & Kurniawan, “Implementasi Confusion Matrix dalam Sistem Klasifikasi Teks Menggunakan Algoritma Naïve Bayes,” J. Teknol. Inf. dan Komput., vol. 9, no. 3, pp. 112–120, 2022.