

# Klasifikasi Penyakit Ginjal Kronis Menggunakan Algoritma Naïve Bayes: Literature Review

Bagus Taufik Hidayat<sup>1\*</sup>, Dani<sup>1</sup>

<sup>1</sup>Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Jl. Raya Puspipetek No. 46, Kel. Buaran, Kec. Serpong, Kota Tangerang Selatan. Banten 15310, Indonesia

Email: [1\\*bagustaufik26@gmail.com](mailto:1*bagustaufik26@gmail.com), [2danicoleon12@gmail.com](mailto:2danicoleon12@gmail.com)

(\* : coressponding author)

**Abstrak** – Penyakit Ginjal Kronis (PGK) merupakan salah satu masalah kesehatan global yang membutuhkan deteksi dini untuk mencegah komplikasi serius. Dalam dunia kesehatan, algoritma Naïve Bayes telah menunjukkan potensi sebagai metode yang efektif untuk klasifikasi data medis, termasuk PGK, karena sifatnya yang sederhana namun akurat dalam menangani data dengan variabel yang saling independen. Studi ini bertujuan untuk melakukan tinjauan literatur terhadap penerapan algoritma Naïve Bayes dalam klasifikasi PGK, dengan fokus pada akurasi, efisiensi, dan keandalan model yang dihasilkan. Penelitian ini menganalisis berbagai studi sebelumnya, termasuk teknik praproses data, fitur penting yang digunakan, serta evaluasi model kinerja berdasarkan parameter seperti accuracy, precision, dan recall. Hasil tinjauan menunjukkan bahwa algoritma Naïve Bayes mempunyai akurasi yang kompetitif untuk mengklasifikasi PGK jika dibandingkan dengan metode lainnya, terutama pada dataset dengan jumlah fitur terbatas. Kesimpulan dari tinjauan ini menyoroti pentingnya pengelolaan data yang optimal dan pemilihan fitur relevan untuk meningkatkan performa algoritma Naïve Bayes. Studi ini diharapkan dapat memberikan panduan bagi peneliti di masa depan untuk mengembangkan sistem deteksi dini PGK berbasis machine learning.

**Kata Kunci:** Penyakit Ginjal Kronis, Naïve Bayes, Klasifikasi, Machine Learning, Studi Literatur

**Abstract** – *Chronic Kidney Disease (CKD) is a significant global health issue that requires early detection to prevent serious complications. In the field of healthcare, the Naïve Bayes algorithm has shown potential as an effective method for classifying medical data, including CKD, due to its simplicity yet accuracy in handling data with independent variables. This study aims to conduct a literature review on the application of the Naïve Bayes algorithm in CKD classification, focusing on the accuracy, efficiency, and reliability of the resulting models. The research analyzes various previous studies, including data preprocessing techniques, important features used, and performance model evaluations based on parameters such as accuracy, precision, and recall. The review findings indicate that the Naïve Bayes algorithm offers competitive accuracy for classifying CKD compared to other methods, especially on datasets with a limited number of features. The conclusion of this review highlights the importance of optimal data management and the selection of relevant features to improve the performance of the Naïve Bayes algorithm. This study is expected to provide guidance for future researchers in developing early detection systems for CKD based on machine learning.*

**Keywords:** *Chronic Kidney Disease, Naïve Bayes, Classification, Machine Learning, Literature Review*

## 1. PENDAHULUAN

Penyakit Ginjal Kronis (PGK) adalah kondisi medis yang ditandai dengan penurunan fungsi ginjal yang terjadi secara progresif dan permanen. Menurut informasi data yang dihimpun dari Organisasi Kesehatan Dunia (WHO), PGK termasuk salah satu penyebab utama tingginya angka mortalitas dan morbiditas global, dengan prevalensinya yang terus mengalami peningkatan setiap tahunnya. Di banyak negara, PGK sering kali tidak terdiagnosis pada tahap awal karena gejalanya yang cenderung tidak spesifik dan berkembang perlahan. Oleh karena itu, PGK sering kali dikenal sebagai penyakit silent disease atau kondisi yang tidak menimbulkan gejala secara jelas. Padahal, deteksi dini PGK sangat penting untuk mencegah komplikasi serius, seperti gagal ginjal stadium akhir yang membutuhkan terapi dialisis atau transplantasi ginjal. Terlambatnya diagnosis sering berdampak pada menurunnya kualitas hidup pasien dan melonjaknya biaya perawatan pasien. Dengan meningkatnya prevalensi PGK, ada kebutuhan mendesak untuk metode yang dapat mendeteksi penyakit ini secara lebih cepat dan akurat pada tahap awal.

Seiring dengan kemajuan teknologi, pendekatan berbasis machine learning semakin populer dalam dunia kesehatan, terutama untuk analisis dan prediksi kondisi medis seperti PGK. Machine learning mempunyai keunggulan dalam memproses dan menganalisis data berukuran besar dengan berbagai variabel yang saling berkaitan. Naïve Bayes adalah salah satu algoritma yang umum

diterapkan dalam klasifikasi medis, terkenal karena kemudahannya dalam implementasi serta efisiensinya dalam menangani data dengan variabel yang dianggap independen. Algoritma ini dapat memberikan hasil yang akurat meskipun dengan dataset yang relatif kecil dan sering digunakan pada aplikasi yang memerlukan keputusan cepat, seperti prediksi penyakit. Keunggulan utama dari Naïve Bayes terletak pada kemampuannya untuk bekerja dengan data yang tidak terstruktur dan hasil yang cepat dengan kebutuhan sumber daya komputasi yang rendah, membuatnya cocok diterapkan dalam sistem prediksi kesehatan berbasis data.

Namun, meskipun Naïve Bayes telah terbukti efektif pada banyak aplikasi, efektivitasnya dalam klasifikasi PGK sangat bergantung pada beberapa faktor, seperti kualitas data yang digunakan, teknik praproses yang diterapkan, serta pemilihan fitur yang tepat. Pada data medis, termasuk PGK, sering kali terdapat tantangan dalam hal missing values, ketidakseimbangan data (imbalanced dataset), dan variabel-variabel yang mungkin saling bergantung. Oleh karena itu, pemahaman yang mendalam tentang bagaimana mempersiapkan dan mengolah data sangat penting untuk meningkatkan performa model. Selain itu, pemilihan fitur yang relevan dan representatif sangat mempengaruhi kemampuan algoritma dalam menghasilkan klasifikasi yang akurat. Dalam konteks PGK, fitur-fitur seperti tekanan darah, kadar gula darah, dan usia, serta riwayat kesehatan keluarga sering digunakan, tetapi pemilihan fitur yang tepat harus didasarkan pada analisis yang teliti terhadap faktor-faktor yang memengaruhi perkembangan PGK.

Penelitian ini bertujuan untuk melakukan kajian literatur yang mendalam mengenai penerapan algoritma Naïve Bayes dalam klasifikasi PGK. Fokus utama dari studi ini adalah untuk mengevaluasi akurasi, efisiensi, serta tantangan dan potensi yang dihadapi oleh algoritma Naïve Bayes dalam konteks pengelolaan data medis, khususnya dalam deteksi dini PGK. Selain itu, studi ini juga akan membahas pentingnya praproses data, pemilihan fitur, serta teknik-teknik yang dapat digunakan untuk meningkatkan kinerja algoritma, termasuk penggunaan teknik resampling untuk menangani ketidakseimbangan data. Dengan tujuan untuk memberikan gambaran yang lebih lengkap mengenai kelebihan dan keterbatasan Naïve Bayes, penelitian ini diharapkan dapat memberikan manfaat terkait pengetahuan yang berguna dalam pengembangan sistem prediksi berbasis machine learning untuk PGK.

## 2. METODOLOGI PENELITIAN

### 2.1 Data Collection

Pengumpulan data dengan proses pencarian dan pemilihan literatur yang relevan dari berbagai sumber yang terpercaya, seperti database akademik dan jurnal peer-reviewed. Pencarian ini dilakukan menggunakan kata kunci tertentu untuk menemukan artikel yang memenuhi kriteria inklusi. Proses pengumpulan data meliputi dua tahap utama: pencarian artikel dan ekstraksi data yang relevan dari jurnal yang telah terpilih, Seperti dari sumber: google scholar, iee explore dan lain sebagainya. Ekstraksi data mencakup informasi tentang metodologi yang digunakan, dataset, teknik praproses data, dan hasil evaluasi kinerja model. Dengan ketentuan jurnal harus terbit lima tahun terakhir untuk memastikan keaktualan data.

### 2.2 Inclusion and Exclusion Criteria

Kriteria inklusi adalah aturan yang digunakan untuk menentukan artikel-artikel yang layak dimasukkan dalam tinjauan literatur. Artikel yang memenuhi kriteria inklusi harus relevan dengan topik penelitian dan sesuai dengan tujuan sistematis review. Misalnya, hanya studi yang menggunakan algoritma Naïve Bayes untuk mengklasifikasi Penyakit Ginjal Kronis (PGK) yang akan dimasukkan. Aturan yang digunakan untuk mengecualikan artikel-artikel yang tidak memenuhi syarat. Artikel yang tidak memenuhi kualitas metodologi, tidak melibatkan Naïve Bayes, atau tidak relevan dengan PGK, akan dikeluarkan dari tinjauan ini.

### 2.3 Study Selection Process

Proses seleksi studi melibatkan dua tahap utama:

- a. Tahap Penyaringan Awal: Dalam tahap ini, artikel disaring berdasarkan judul dan abstrak untuk memastikan kesesuaiannya dengan topik dan kriteria inklusi. Artikel yang tidak relevan atau tidak memenuhi kriteria akan dihapus.

- b. Tahap Penyaringan Lengkap: Pada tahap ini, artikel yang lolos dari penyaringan awal dibaca secara penuh untuk memverifikasi kelayakannya dan memastikan bahwa mereka memenuhi kriteria inklusi. Setelah tahap ini, artikel yang memenuhi kriteria akan dimasukkan ke dalam analisis lebih lanjut.

### 3. ANALISA DAN PEMBAHASAN

#### 3.1 Tabel Analisa

Pada bagian ini berisi temuan, hasil serta pembahasan topik penelitian.

**Tabel 1.** Hasil Rangkuman Artikel yang Dianalisa

No	Nama Peneliti dan Tahun	Metode yang dibahas	Tujuan Penelitiannya	Hasil yang didapat
1	Ena Tasia, Qurotul A'yuniyah, Pangeran Faddilah Pratama, Nanda Nazira, Mustakim, Jeni Adhiva, M. Ridho Anugrah	Implementasi Algoritma Naïve Bayes Classifier (NBC) untuk Klasifikasi Penyakit Ginjal Kronik	Penelitian ini bertujuan untuk mengklasifikasikan data Penyakit Ginjal Kronik (PGK) menggunakan algoritma Naïve Bayes, serta mengevaluasi akurasi, recall, precision, dan AUC untuk menilai kinerja model dalam mendeteksi PGK.	Klasifikasi menggunakan algoritma Naïve Bayes (NBC) melalui tools RapidMiner menghasilkan akurasi tinggi sebesar 96,43%, menunjukkan efektivitas algoritma dalam mengklasifikasikan data PGK.
2	Adhi Indra Irawan, Ristu Saptono, Afrizal Doewes	Implementasi On Of Naive Bayes Classifier Method And Adaboost Algorithm For Prediction Of Chronic Kidney Disease	Penelitian ini bertujuan untuk mengaplikasikan metode Naïve Bayes dan algoritma AdaBoost dalam proses klasifikasi penyakit ginjal kronis (CKD), serta mengevaluasi kinerja kedua metode dalam menangani nilai yang hilang pada dataset medis.	Metode Naïve Bayes efektif mengklasifikasikan data CKD meski terdapat banyak missing value, dengan akurasi 0,95. Dengan menggunakan AdaBoost, akurasi meningkat menjadi 0,98 dan F1 score tercatat sebesar 0,984, baik pada dataset asli maupun setelah penggantian missing value.
3	Firdhani Novrizal, M.Iwan Wahyudin, Ratih Titi Komalasari	Expert System to Diagnose Kidney Failure with Web-Based Naïve Bayes Method	Penelitian ini berfokus pada pengembangan sistem pakar berbasis metode Naïve Bayes untuk memberikan bantuan kepada masyarakat dan dokter dalam mengantisipasi serta mendiagnosis gagal ginjal, dengan akurasi 81,67% dari 60 data uji.	Metode Naïve Bayes dalam sistem pakar ini menunjukkan efektivitas yang tinggi dalam mendiagnosis gagal ginjal, membantu dokter dalam penentuan diagnosis awal, dan memberikan hasil persentase dengan akurasi 81,67%, memudahkan masyarakat mengenali gejala sejak dini.
4	Nandi Prabu Nugraha, Putri Salma Ningayu, Syauqi Zalfha Daffa, Rafian Azim,	Perbandingan Akurasi Metode Naïve Bayes dan Metode KNN untuk	Penelitian ini bertujuan membandingkan akurasi metode Klasifikasi Naïve Bayes dan K -Nearest Neighbors (KNN) dalam mendiagnosis penyakit	Klasifikasi Naïve Bayes dan KNN keduanya memiliki prediksi tinggi, dengan Naïve Bayes sedikit lebih rendah dalam akurasi. Naïve Bayes lebih mudah

		Memprediksi Gagal Ginjal Kronis.	gagal ginjal kronis, untuk menentukan metode yang lebih tepat digunakan pada data besar dan tidak lengkap.	diterapkan, sementara KNN membutuhkan usaha ekstra untuk menentukan jumlah neighbors (K). Disarankan untuk menguji kedua metode dengan dataset bervariasi untuk menentukan yang lebih akurat.
5	Alif Nur Fadilah, Alifah Dafa Iftinani, Ali Nur Ikhsan,	Performance Comparison of Decision Tree J48, CART and Naïve Bayes Algorithms for Predicting Chronic Kidney Disease	Tujuan penelitian ini adalah untuk membandingkan kinerja Naive Bayes, CART, dan algoritma Decision Tree J48 dalam memprediksi gagal ginjal kronis, serta untuk menentukan algoritma yang paling akurat di antara ketiganya sebagai dasar untuk pengembangan program prediksi penyakit ginjal kronis.	Algoritma CART memiliki akurasi tertinggi (97,25%) dalam memprediksi penyakit ginjal kronis, mengungguli J48 (96,5%) dan Naïve Bayes (93,5%). Algoritma CART dapat digunakan untuk pengembangan program prediksi, dan penelitian selanjutnya disarankan untuk mengoptimalkan algoritma lain dengan teknik seperti PSO atau Algoritma Genetik.

### 3.2 Pembahasan

Penelitian-penelitian terkait klasifikasi penyakit ginjal menggunakan berbagai algoritma machine learning, terutama Naïve Bayes (NBC), menunjukkan beragam hasil yang menggambarkan efektivitas algoritma dalam mengklasifikasikan data medis. Pada penelitian pertama, penggunaan Naïve Bayes dengan tools RapidMiner untuk mengklasifikasikan data Penyakit Ginjal Kronis (PGK) menghasilkan akurasi tinggi sebesar 96,43%. Naïve Bayes terbukti efektif dalam klasifikasi data PGK, yang dipengaruhi oleh kesederhanaan serta kemampuannya dalam memproses data yang besar.

Penelitian kedua juga mengonfirmasi efektivitas Naïve Bayes dalam mengklasifikasikan data Penyakit Ginjal Kronis (CKD), meskipun terdapat banyak missing values. Akurasi awal yang dicapai adalah 0,95, namun setelah penerapan AdaBoost untuk meningkatkan model, akurasi meningkat menjadi 0,98 dengan F1 score 0,984, menunjukkan peningkatan signifikan dalam kinerja model. Ini menunjukkan bahwa teknik ensemble seperti AdaBoost dapat meningkatkan performa Naïve Bayes, terutama dalam menangani data yang memiliki banyak nilai hilang. Pada penelitian ketiga, sistem pakar berbasis Naïve Bayes digunakan untuk mendiagnosis gagal ginjal dengan akurasi 81,67%. Sistem ini memudahkan dokter dalam melakukan diagnosis awal dan membantu masyarakat mengenali gejala penyakit ginjal sejak dini. Meskipun akurasinya bisa lebih tinggi, sistem ini tetap menawarkan solusi yang berguna dalam aplikasi medis praktis.

Penelitian keempat melakukan perbandingan antara Naive Bayes dan algoritma K-Nearest Neighbors (KNN) untuk mengklasifikasikan penyakit ginjal. Kedua algoritma menunjukkan hasil prediksi yang tinggi, meski demikian Naive Bayes menghasilkan akurasi yang sedikit lebih rendah dibandingkan dengan menggunakan KNN. Naïve Bayes lebih mudah diterapkan karena tidak memerlukan parameter tambahan seperti penentuan jumlah neighbors (K) yang diperlukan oleh KNN. Penelitian ini menyarankan untuk menguji kedua algoritma dengan dataset yang lebih beragam untuk menentukan metode yang lebih akurat dalam konteks yang berbeda.

Pada penelitian terakhir, algoritma Classification and Regression Tree (CART) menunjukkan akurasi tertinggi (97,25%) dibandingkan dengan J48 (96,5%) dan Naïve Bayes (93,5%) dalam memprediksi penyakit ginjal kronis. Keunggulan CART terletak pada kemampuannya untuk menangani data yang lebih kompleks dan menghasilkan model yang lebih mudah dipahami.

Penelitian ini juga menyarankan untuk mengoptimalkan algoritma lain, seperti J48 atau Naïve Bayes, menggunakan teknik optimasi seperti Particle Swarm Optimization (PSO) atau Algoritma Genetik untuk meningkatkan performa.

Secara umum, hasil penelitian ini mengindikasikan bahwa Naive Bayes efektif untuk mengklasifikasikan penyakit ginjal, walaupun pada beberapa kondisi hasilnya berbeda, peningkatan performa dapat dicapai dengan teknik ensemble atau algoritma lain seperti CART. Hal ini menggarisbawahi pentingnya memilih algoritma yang tepat, yang sangat dipengaruhi oleh sifat dataset dan teknik optimasi yang digunakan.

#### 4. KESIMPULAN

Kesimpulan dari jurnal ini menunjukkan algoritma Naive Bayes (NBC) efektif dalam mengklasifikasikan data penyakit gagal ginjal kronis (PGK), dengan akurasi yang sangat tinggi (hingga 96,43%) menggunakan tools RapidMiner. Algoritma ini juga mampu menangani missing value dengan baik, meskipun penggantian missing value dan penggunaan AdaBoost dapat meningkatkan akurasi lebih lanjut, mencapai 98% dan F1-score 0,984. Sistem pakar berbasis Naive Bayes memberikan hasil yang memadai (81,67%) dalam membantu diagnosis gagal ginjal, memberikan manfaat bagi deteksi dini oleh masyarakat dan tenaga medis.

Selain itu, meskipun akurasi Klasifikasi Naïve Bayes sedikit lebih rendah dibandingkan dengan KNN, Naïve Bayes lebih mudah diterapkan, sedangkan KNN memerlukan penentuan jumlah neighbors (K) yang lebih rumit. Dalam perbandingan dengan algoritma lain, seperti CART yang mencapai akurasi tertinggi (97,25%), Naïve Bayes terbukti efektif meskipun ada algoritma lain yang memberikan hasil lebih baik dalam beberapa kasus. Oleh karena itu, Peneliti selanjutnya dapat mengeksplorasi lebih jauh dengan untuk mengoptimalkan algoritma dengan menggunakan teknik seperti PSO atau Algoritma Genetik, guna meningkatkan prediksi penyakit ginjal kronis secara lebih optimal.

#### REFERENCES

- A'yuniyah, Q., Tasia, E., Nazira, N., Pratama, P. F., Anugrah, M. R., Adhiva, J., & Mustakim, M. (2022). Pemanfaatan Algoritma Naïve Bayes Classifier (NBC) dalam Proses Klasifikasi Penyakit Ginjal Kronik. *Jurnal Sistem Komputer dan Informatika (JSON)*, 4(1), 72-76.
- Wulandari, V., Sari, W. J., Alfian, Z., Legito, L., & Arifianto, T. (2024). Penerapan Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor untuk Penentuan Klasifikasi Penyakit Ginjal Kronik. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(2), 710-718.
- Novrizal, F., Wahyuddin, M. I., & Komalasari, R. T. (2020). Sistem Pakar Berbasis Web Menggunakan Metode Naïve Bayes untuk Diagnosis Gagal Ginjal. *Jurnal Mantik*, 3(4), 466-473.
- Nugraha, N. P., Azim, R., Daffa, S. Z., & Ningayu, P. S. (2023). Analisis Perbandingan Akurasi Metode Naïve Bayes dan K-Nearest Neighbor pada Prediksi Gagal Ginjal Kronis. *Jurnal Rekayasa Elektro Sriwijaya*, 5(1), 1-10.
- Ikhsan, A. N., Fadilah, A. N., & Iftinani, A. D. (2024). Performance Comparison of Decision Tree J48, CART, and Naïve Bayes Algorithms for Predicting Chronic Kidney Disease. *Indonesian Journal of Artificial Intelligence and Data Mining*, 7(1), 64-70.
- Ezaputra, A. R., Hidayat, E. R., Darmawati, L. S. N., & Ikasari, I. H. (2024). Tinjauan Literatur: Inovasi dalam Pembelajaran Pemrograman dan Sistem Informasi. *JRIIN: Jurnal Penelitian Informatika dan Inovasi*, 2(4), 607-614.
- Anggraini, Y., Indra, M., Khoirusofi, M., Azis, I. N., & Rosyani, P. (2023). Tinjauan Literatur Sistem Pakar untuk Diagnosa Penyakit Gigi dengan Menggunakan Pendekatan Forward Chaining. *BINER: Jurnal Ilmu Komputer, Teknik dan Multimedia*, 1(1), 1-7.
- Kausar, A., Irawan, A., & Fernando, I. (2023). Implementasi Algoritma Naïve Bayes untuk Penilaian Kinerja Dosen. *PROSISKO: Jurnal Pengembangan Riset dan Observasi Sistem Komputer*, 10(2), 117-127.



- Komala, L., Pamungkas, I. B., & Rodiyana, N. (2023) Analisis Pengaruh Teknologi Informasi dan Motivasi terhadap Kinerja: Tinjauan Berdasarkan Literatur. *Scientific Journal of Reflection: Economic, Accounting, Management and Business*, 6(3), 716-724.
- Ishlah, F. M., Khatami, I., Rizqi, M. F. N., Marcus, Y. M., & Rosyani, P. (2023). Studi Literatur Mengenai Sistem Pakar dengan Pendekatan Forward Chaining. *Journal of Research and Publication Innovation*, 1(3), 574-578.
- Husna, N. C. (2012). Gagal Ginjal Kronis dan Penanganannya: Sebuah Literatur Review. *FIKkeS*, 3(2).
- Esthi, S. W., Hanifsyah, D. M., Gabe, A., & Rosyani, P. (2024) Tinjauan Literatur Sistem Pakar Menggunakan Metode Case Based Reasoning (CBR) dalam Diagnosa Penyakit Usus. *LOGIC: Jurnal Ilmu Komputer dan Pendidikan*, 2(2), 409-412.
- Oktalia, C., & Zakaria, H. (2024). Perbandingan Kinerja Algoritma C4.5 dan Naïve Bayes dalam Menganalisis Ulasan Pengguna Layanan JNE pada Aplikasi Google Playstore. *BINER: Jurnal Ilmu Komputer, Teknik dan Multimedia*, 2(3), 271-284.
- Hendry, M., & Djaksana, Y. M. (2024). Penerapan Metode Decision Tree C4.5 dan Naïve Bayes pada Klasifikasi Karakteristik Kepribadian Manusia. *OKTAL: Jurnal Ilmu Komputer dan Sains*, 3(03), 771-777.
- Alfian, Z. (2022). Prediksi Pembelian Stok Barang dengan Menggunakan Algoritma Naïve Bayes: Studi Kasus pada CV. Kurnia Jaya. *Scientia Sacra: Jurnal Sains, Teknologi dan Masyarakat*, 2(3), 451-454.