JURNAL ILMIAH SISTEM INFORMASI (JUSI)
VOL 4, No. 1, Januari 2025, pp. 180 - 190
p-ISSN : 2803-1507 (media cetak) | e-ISSN : 2803-1531 (media online)
Link Publikasi: https://ejurnal.provisi.ac.id/index.php/JUISI/issue/view/126

Chronic Kidney Disease Prediction Model Using Naïve Bayes (Case Study: Jayapura City)

Grace Adelin Rumbairusy*[1], Manda[2], Yulan Nanda Sandira Payungallo[3], Heru Sutejo[4]
[1,2,3,4] Program Studi Sistem Informasi, Universitas Sepuluh Nopember Papua
Jl. Ardipura II No. 22, 99222, Jayapura, Papua, Indonesia
e-mail: rumbairusy30@gmail.com*[1], heru.sutejo01@gmail.com[4]

ARTICLE INFO

Article History:

Received: 02-10-2025
Revised: 21-11-2025
Accepted: 03-12-2025
Published: 30-12-2025

ABSTRACT

Chronic Kidney Disease (CKD) remains one of the most critical global health issues due to its progressive nature and the absence of early symptoms that often lead to late diagnosis. In Jayapura City, the number of CKD cases continues to increase, driven by factors such as hypertension, diabetes, and limited access to early diagnostic services. This study aims to develop a predictive model for CKD using the Naïve Bayes algorithm and to analyze the relationship between clinical variables that contribute to CKD among patients in Jayapura. The research adopts the Cross-Industry Standard Process for Data Mining (CRISP-DM), which includes business understanding, data understanding, data preparation, modeling, evaluation, and deployment. A clinical dataset consisting of 500 patient records and 13 medical attributes—such as blood pressure, glucose level, serum creatinine, hemoglobin, albumin, and urine characteristics—was used. The dataset had been pre-cleaned before processing, ensuring it was ready for direct modeling in RapidMiner. The Naïve Bayes model achieved an accuracy of 94.40%, with substantial precision and recall values across both CKD and non-CKD classes. These findings indicate that Naïve Bayes is highly effective in identifying CKD patterns within local patient data. The main contribution of this study lies in the use of real clinical data specific to Jayapura, which provides insights into regional CKD characteristics and offers a reliable predictive tool to support early screening. The implications of this research highlight the potential integration of the model into clinical decision-support systems and community health monitoring applications, enabling earlier interventions and improved patient outcomes.

Keywords: *Chronic Kidney Disease, Naïve Bayes, RapidMiner, Medical Prediction, Data Mining.*

Abstrak

Penyakit Ginjal Kronis (PGK) merupakan salah satu masalah kesehatan yang kritis karena bersifat progresif dan sering tidak menunjukkan gejala pada tahap awal, sehingga banyak pasien terdiagnosis pada stadium lanjut. Di Kota Jayapura, jumlah kasus PGK terus meningkat akibat hipertensi, diabetes, serta keterbatasan akses layanan deteksi dini. Penelitian ini bertujuan membangun model prediksi PGK menggunakan algoritma *Naïve Bayes* serta menganalisis keterkaitan variabel klinis yang berpengaruh terhadap PGK pada pasien di

DOI: https://doi.org/10.51903/k47t6677

p-ISSN : 2803-1507 e-ISSN : 2803-1531

Jayapura. Penelitian mengikuti kerangka kerja *Cross-Industry Standard Process for Data Mining (CRISP-DM)*, meliputi pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan *deployment*. Dataset yang digunakan terdiri dari 500 data pasien dengan 13 atribut medis, termasuk tekanan darah, glukosa darah, kreatinin serum, *hemoglobin, albumin*, dan kondisi urin. Seluruh data telah melalui tahap pembersihan sebelum pemodelan sehingga tidak memerlukan preprocessing lanjutan. Pemodelan dilakukan menggunakan *RapidMiner*, dan algoritma *Naïve Bayes* menghasilkan akurasi sebesar 94.40% dengan nilai *precision* dan *recall* tinggi pada kedua kelas PGK dan non-PGK. Hasil ini menunjukkan bahwa *Naïve Bayes* efektif dalam mengidentifikasi pola PGK pada data klinis lokal. Kontribusi utama penelitian ini adalah pemanfaatan data nyata dari pasien Kota Jayapura, sehingga menghasilkan model prediksi yang relevan secara regional serta memberikan pemahaman baru mengenai faktor medis yang dominan. Implikasi penelitian ini mencakup potensi integrasi model ke dalam sistem pendukung keputusan klinis maupun aplikasi monitoring kesehatan untuk mendukung deteksi dini PGK dan meningkatkan kualitas layanan kesehatan.

Kata Kunci: *Penyakit Ginjal Kronis, Naïve Bayes, RapidMiner, Prediksi Medis, Data Mining.*

1. PENDAHULUAN

Penyakit ginjal kronis merupakan gangguan fungsi ginjal yang berlangsung secara progresif dan tidak dapat dipulihkan, sehingga tubuh tidak mampu menjaga keseimbangan metabolisme, cairan, dan elektrolit. Jika terus berlanjut, kondisi ini dapat berkembang menjadi gagal ginjal tahap akhir, yaitu saat ginjal tidak lagi mampu membuang zat-zat sisa dari tubuh. Pada tahap ini, pasien memerlukan tindakan lanjutan seperti dialisis (cuci darah) atau transplantasi ginjal untuk mempertahankan kehidupannya [1]. *World Health Organization (WHO)* mencatat PGK sebagai penyebab kematian ke-10 di dunia, sedangkan Kementerian Kesehatan Republik Indonesia melaporkan bahwa prevalensi PGK meningkat dalam lima tahun terakhir akibat pola hidup tidak sehat, hipertensi, dan diabetes sebagai faktor risiko dominan [2], [3].

Di Kota Jayapura, fenomena serupa tampak pada data pasien yang menunjukkan peningkatan jumlah penderita gangguan fungsi ginjal, yang tercermin dari tingginya kadar kreatinin serum, urea darah, tekanan darah yang tidak terkendali, serta gejala klinis lain yang sering tidak terdeteksi pada tahap awal. Keterbatasan akses kesehatan di beberapa distrik di Jayapura menyebabkan pasien datang pada stadium lanjut, sehingga deteksi dini dengan metode prediktif berbasis data menjadi sangat penting untuk membantu klinisi dalam pengambilan keputusan. Kondisi ini menunjukkan bahwa variasi gejala yang muncul pada pasien sering kali berkaitan dengan perbedaan gaya hidup dan akses layanan kesehatan. Data tersebut juga memperlihatkan bahwa perubahan pola fungsi ginjal pada masyarakat Jayapura cenderung berkembang secara perlahan sehingga sulit dikenali tanpa pemeriksaan klinis rutin.

Berbagai penelitian terkait prediksi PGK telah memanfaatkan algoritma *machine learning* seperti *Naïve Bayes*, *K-Nearest Neighbour, Decision Tree,* dan Support *Vector Machine* [4], [5], [6], [7]. *Naïve Bayes* terbukti memiliki kinerja yang kompetitif pada dataset medis yang memiliki jumlah sampel terbatas dan fitur klinis yang beragam [8], [9], [10], serta efektif dalam mengklasifikasikan data dengan asumsi independensi antar variabel [11]. Beberapa penelitian sebelumnya melakukan klasifikasi penyakit ginjal kronis *(Chronic Kidney Disease*/CKD) dengan Algoritma *Naïve Bayes* berhasil mencapai akurasi sebesar 95.4% dalam memprediksi penyakit ginjal kronis [12]. Penelitian lain menyimpulkan bahwa algoritma *Naïve Bayes* sangat efektif untuk mendiagnosis penyakit ginjal, dengan akurasinya mencapai 99.24% [13].

Selain itu dengan eksperimen yang di lakukan [14] menunjukkan bahwa *Naïve Bayes* memberikan performa terbaik, dengan akurasi tertinggi sebesar 94,21%. Sementara itu, [15]. memperoleh akurasi sebesar 90% pada penelitian mereka menggunakan *Naïve Bayes Classification* berdasarkan hasil confusion matrix. Penelitian lain Penggunaan algoritma *Naive Bayes* dalam memprediksi penyakit *chronic kidney disease* (CKD) terbukti memberikan kinerja yang baik dan menawarkan berbagai manfaat di bidang kesehatan. Dengan memanfaatkan kumpulan data klinis dan laboratorium yang lengkap sebagai atribut, serta didukung oleh metode seleksi fitur yang optimal dan potensi penerapan ensemble, algoritma *Naive Bayes* mampu menghasilkan prediksi CKD yang tepat dan efisien [16].

Namun, sebagian besar penelitian tersebut menggunakan dataset UCI atau data dari rumah sakit besar di Pulau Jawa, sehingga belum merepresentasikan kondisi klinis pasien di Papua. Belum ada penelitian yang secara khusus menganalisis pola keterkaitan variabel medis seperti tekanan darah, glukosa darah, kreatinin serum, urea darah, albumin serum, hemoglobin, elektrolit (natrium dan kalium), kondisi urine, dan faktor risiko komorbid (hipertensi, diabetes, penyakit jantung, riwayat merokok) pada populasi pasien di Kota

p-ISSN : 2803-1507 e-ISSN : 2803-1531

Jayapura. Ketiadaan kajian berbasis data lokal ini membentuk research gap penting terkait potensi perbedaan karakteristik klinis maupun faktor prediktif PGK pada masyarakat Papua. Kondisi ini menunjukkan bahwa karakteristik klinis penduduk Papua memerlukan pemetaan yang lebih spesifik agar analisis prediksi dapat menggambarkan situasi kesehatan secara lebih akurat.

Berdasarkan kesenjangan tersebut, penelitian ini bertujuan untuk mengidentifikasi pola dan keterkaitan antara berbagai variabel medis dalam dataset Kota Jayapura, terutama tekanan darah, glukosa darah, kreatinin serum, urea darah, albumin serum, serta komorbid seperti hipertensi dan diabetes yang berkontribusi terhadap risiko terjadinya PGK. Selain itu, penelitian ini juga bertujuan membangun model prediksi PGK menggunakan metode *Naïve Bayes* dan mengevaluasi performanya pada data pasien lokal Jayapura. Kontribusi atau *novelty* penelitian ini terletak pada pemanfaatan dataset klinis nyata dari Kota Jayapura, sehingga menghasilkan model prediksi PGK yang lebih relevan secara regional, sekaligus memberikan pemahaman baru mengenai faktor-faktor klinis yang paling berpengaruh dalam mendeteksi risiko PGK pada populasi Papua. Model prediksi yang dikembangkan melalui data lokal ini juga diharapkan dapat memberikan gambaran yang lebih rinci mengenai variasi respons klinis yang muncul pada pasien dari wilayah tersebut.

2. TINJAUAN PUSTAKA

2.1. *Data Mining*

Perkembangan penelitian terkait pemanfaatan data klinis mendorong penggunaan teknik *data mining* dalam analisis kesehatan. Penelitian pada [17] menunjukkan bahwa metode *data mining* mampu mengekstraksi pola dari data laboratorium dan riwayat penyakit yang memiliki struktur kompleks. Studi tersebut menjelaskan bahwa sebagian besar data medis masih berupa data mentah yang memerlukan proses analitik sistematis agar menghasilkan informasi yang bermanfaat. Teknik *data mining* digunakan secara luas dalam penelitian prediksi PGK karena proses ini dapat membantu mengatasi masalah variabilitas data, *noise*, dan ketidaklengkapan informasi.

2.2. *Naïve Bayes*

Penelitian pada [18] memberikan landasan mengenai struktur kerja Naïve Bayes sebagai algoritma klasifikasi berbasis Teorema Bayes beserta asumsi independensi antar atributnya. Penelitian tersebut menunjukkan bahwa meskipun asumsi ini tidak selalu terpenuhi pada data medis, Naïve Bayes tetap memberikan *performance* yang kompetitif. Kajian lain [19] menunjukkan bahwa algoritma ini dapat bekerja efektif pada *dataset* klinis yang berukuran terbatas dan memiliki keragaman variabel. Hasil dari penelitian tersebut menjelaskan bahwa Naïve Bayes mampu menangani data laboratorium yang sering mengandung *noise* dan variasi antar sampel.

2.3. *Rapid Miner*

Perkembangan alat bantu analisis data juga menjadi perhatian berbagai penelitian. Studi pada [20], [21] mengevaluasi penggunaan RapidMiner sebagai platform *data mining* yang menyediakan beragam operator untuk *preprocessing*, pemodelan, dan evaluasi. Kedua studi tersebut menjelaskan bahwa RapidMiner memungkinkan implementasi algoritma seperti Naïve Bayes, KNN, dan Decision Tree secara efisien melalui antarmuka visual yang mudah digunakan. Penggunaan platform ini menjadi relevan dalam penelitian yang membutuhkan proses eksperimen yang terstruktur pada data PGK.

2.4. *Penyakit Ginjal Kronis*

Kajian mengenai penyakit ginjal kronis (PGK) banyak dilakukan karena kondisi ini merupakan salah satu masalah kesehatan global. Penelitian pada [22] menjelaskan bahwa PGK merupakan penurunan fungsi ginjal yang berlangsung minimal tiga bulan dan sering berhubungan dengan komorbid seperti diabetes melitus maupun hipertensi. Studi lain [23] membahas perkembangan PGK dari stadium awal hingga stadium terminal yang memerlukan terapi pengganti ginjal. Penelitian pada [24] menyajikan variabel klinis yang digunakan dalam penilaian fungsi ginjal, seperti kreatinin serum, urea darah, albumin serum, hemoglobin, tekanan darah, elektrolit, dan kondisi urine. Ketiga studi tersebut menjadi dasar dalam penentuan atribut yang digunakan dalam model prediksi PGK pada penelitian ini.

2.5. *Penelitian Sebelumnya*

Kajian mengenai penerapan metode klasifikasi dalam bidang kesehatan telah menunjukkan perkembangan yang signifikan dalam beberapa tahun terakhir. Penelitian pada [25] menyajikan analisis mengenai efektivitas

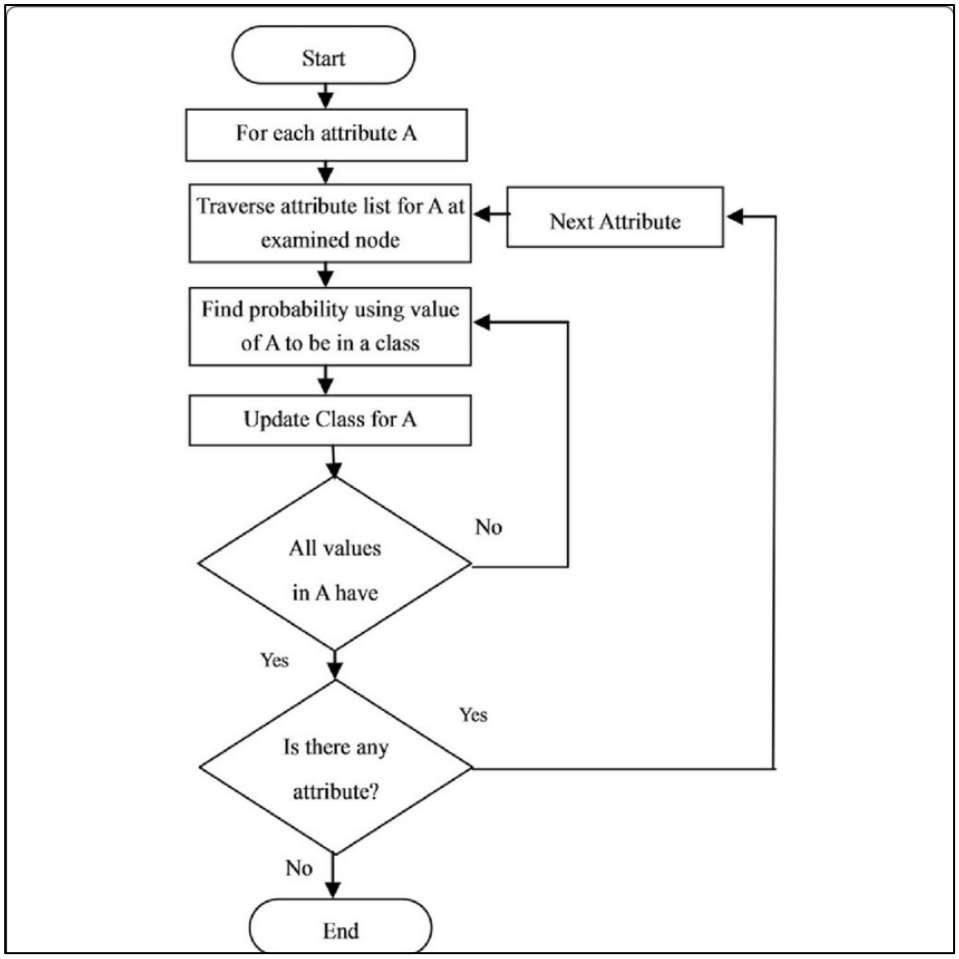
p-ISSN : 2803-1507 e-ISSN : 2803-1531

algoritma Naïve Bayes pada berbagai permasalahan klasifikasi yang memiliki karakteristik berbeda. Studi tersebut menemukan bahwa asumsi independensi antar fitur menghasilkan proses komputasi yang sederhana dan efisien. Penelitian lain [26], [27], [28] menilai penggunaan Naïve Bayes pada domain klasifikasi dan prediksi penyakit, serta melaporkan kemampuan algoritma ini untuk bekerja stabil pada *dataset* berukuran kecil maupun besar.

Beberapa penelitian mengenai prediksi penyakit ginjal kronis (PGK) menggunakan *dataset* yang memiliki tingkat kebersihan tinggi seperti *UCI*, sehingga akurasi yang dilaporkan berkisar 90 hingga 99 persen. Nilai tersebut berkaitan erat dengan kualitas *dataset* yang terstruktur, ukuran sampel yang relatif kecil, dan proporsi kelas yang seimbang. Kondisi tersebut berbeda dengan data klinis yang dihasilkan fasilitas pelayanan kesehatan di Indonesia yang lebih beragam, mengandung ketidakseimbangan kelas, serta memiliki inkonsistensi pencatatan. Penjelasan pada [25], [26], [27], [28] menunjukkan bahwa karakteristik *dataset* sangat memengaruhi performa algoritma, termasuk Naïve Bayes. Penggunaan *dataset* lokal, seperti yang digunakan dalam penelitian ini, memberikan gambaran yang lebih relevan mengenai kinerja model prediksi PGK berdasarkan kondisi data klinis yang nyata.

3. METODOLOGI PENELITIAN

Berikut tahapan alur kerja algoritma *Naïve Bayes* yang digunakan dalam penelitian ini. Flowchart ini menunjukkan urutan proses mulai dari pemilihan setiap atribut, penelusuran nilai atribut pada dataset, perhitungan probabilitas kondisional, hingga pembaruan skor kelas untuk menentukan kemungkinan seorang pasien termasuk kategori CKD atau non-CKD. Melalui tahapan yang sistematis ini, model *Naïve Bayes* dapat mengolah seluruh atribut kesehatan secara independen dan menghasilkan prediksi yang akurat berdasarkan data medis yang tersedia. Berikut *flowchart Naïve Bayes* dapat dilihat pada Gambar 1.



Gambar 1. Flowchart Naive Bayes

p-ISSN : 2803-1507 e-ISSN : 2803-1531

Flowchart tersebut menjelaskan mekanisme dasar algoritma *Naïve Bayes* dalam memproses setiap atribut secara independen untuk menghitung probabilitas suatu data termasuk ke dalam kelas tertentu. Tahapan ini menggambarkan bagaimana model membangun keputusan klasifikasi berdasarkan distribusi nilai atribut yang dianalisis satu per satu. Analisis pada tahap ini memberikan gambaran lebih rinci mengenai bagaimana setiap atribut memberikan kontribusi terhadap proses pengambilan keputusan klasifikasi. Penjelasan tahapan *flowchart* antara lain:

3.1. *Start (Persiapan Studi dan Pengambilan Data)*

Proses dimulai dengan menentukan tujuan penelitian, yaitu memprediksi apakah pasien termasuk kategori CKD atau non-CKD. Data klinis pasien meliputi tekanan darah, kreatinin serum, hemoglobin, albumin serum, elektrolit, kondisi urine, serta riwayat hipertensi dan diabetes dipersiapkan sebagai input, termasuk pengecekan kelengkapan dan kebersihan data. Langkah ini memastikan bahwa setiap variabel klinis telah ditetapkan secara tepat sebelum diproses oleh model prediksi. Validitas data pada tahap awal juga berperan penting dalam menjaga keakuratan analisis yang dilakukan.

3.2. *Penentuan Atribut dan Kelas Target*

Atribut target ditetapkan sebagai kelas (CKD / non-CKD), sementara seluruh variabel klinis berfungsi sebagai prediktor. Pada tahap ini ditentukan apakah atribut bersifat numerik atau kategorikal, karena masing-masing diproses dengan pendekatan probabilistik yang berbeda. Klasifikasi jenis atribut ini memberikan arahan yang jelas mengenai metode perhitungan peluang yang akan digunakan dalam proses pemodelan. Pemahaman terhadap karakteristik atribut tersebut membantu meningkatkan kualitas hasil prediksi yang dihasilkan algoritma.

3.3. *Perhitungan Distribusi Setiap Atribut (Per Atribut A)*

Pada tahap ini, model menghitung distribusi setiap atribut, baik kategorikal maupun numerik, untuk masing-masing kelas. Untuk atribut kategorikal seperti *hypertension*, *diabetes*, dan *protein urine*, probabilitas kondisional diperoleh dengan menghitung frekuensi kemunculan setiap kategori pada tiap kelas. Probabilitas tersebut dihitung menggunakan persamaan berikut (1).

$$p(A = v \mid Class) = \frac{Jumlah\ record\ dengan\ A = v\ Pada\ kelas}{Jumlah\ record\ di\ kelas} \qquad (1)$$

Pendekatan ini memungkinkan model menangkap kecenderungan kemunculan kategori tertentu dalam masing-masing kelas. Untuk menghindari probabilitas nol pada kategori yang jarang muncul, Laplace Smoothing digunakan sebagai metode penyesuaian.

Untuk atribut numerik seperti *serum creatinine*, *blood urea*, *sodium*, dan *potassium*, nilai mean (μ) dan variance (σ^2) dihitung secara terpisah untuk setiap kelas. Atribut numerik diasumsikan mengikuti distribusi Gaussian sehingga probabilitasnya dimodelkan menggunakan persamaan (2).

$$p(x \mid Class)\frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad (2)$$

Formulasi ini memberikan estimasi probabilistik yang lebih representatif untuk data kontinu. Parameter-parameter tersebut kemudian menjadi dasar dalam penyusunan model prediksi kelas.

3.4. *Penggabungan Probabilitas untuk Prediksi Kelas*

Setelah seluruh probabilitas kondisional dihitung, nilai-nilai tersebut digabungkan dengan probabilitas prior dari masing-masing kelas untuk menghasilkan probabilitas posterior. Proses penggabungan dilakukan menggunakan persamaan (3).

$$P(\ Class \mid A1, A2, \dots, An\ ) \propto P(Class) \times \prod_{i=1}^{n} p(\ A_i \mid Class\ ) \qquad (3)$$

p-ISSN : 2803-1507 e-ISSN : 2803-1531

Nilai posterior yang dihasilkan merepresentasikan tingkat keyakinan model terhadap setiap kelas. Kelas dengan nilai posterior tertinggi dipilih sebagai hasil prediksi akhir, yaitu CKD atau non-CKD.

3.5. *Evaluasi Model*

Model kemudian diuji pada data testing untuk menghitung performa menggunakan metrik seperti akurasi, *precision, recall,* dan *sensitivity*. Tahap ini bertujuan memastikan bahwa model *Naïve Bayes* dapat melakukan prediksi PGK secara konsisten pada data klinis Kota Jayapura. Pengujian ini memberikan gambaran mengenai kemampuan model dalam membedakan pasien yang tergolong CKD dan non-CKD berdasarkan pola atribut klinis yang tersedia. Informasi tersebut juga membantu memvalidasi stabilitas model ketika diterapkan pada variasi data yang berbeda.

4. HASIL DAN PEMBAHASAN

4.1. *Bussines Understanding*

Data Mining ini bertujuan untuk membangun model prediksi penyakit ginjal kronis (*Chronic Kidney Disease* / CKD) berdasarkan data rekam medis pasien. Dengan meningkatnya kasus CKD yang kerap terlambat didiagnosis, diperlukan sebuah sistem pendukung keputusan yang mampu mengidentifikasi kondisi pasien secara cepat berdasarkan parameter klinis seperti tekanan darah, glukosa darah, kreatinin serum, urea darah, dan lain-lain. Dengan menggunakan algoritma *Naive Bayes*, model dapat memberikan prediksi awal apakah seorang pasien termasuk kategori CKD (1) atau non-CKD (0) sehingga tenaga medis dapat melakukan tindak lanjut lebih cepat.

4.2. *Data Understanding*

Dataset yang digunakan terdiri dari 500 data pasien dengan 13 atribut, yang meliputi variabel numerik dan kategorikal. Setiap atribut memiliki karakteristik yang berbeda sehingga diperlukan pemahaman awal untuk melihat distribusi serta perannya dalam analisis. Informasi tersebut membantu peneliti dalam mengenali pola awal yang muncul pada data sebelum memasuki tahap pemodelan. Atribut kunci yang digunakan antara lain dapat dilihat pada Tabel 1.

Tabel 1. Data Atribut

No	Atribut	Deskripsi
1	usia	umur pasien
2	jenis_kelamin	M/F
3	tekanan_darah	tekanan darah mmHg
4	glukosa_darah	gula darah mg/dL
5	kreatinin_serum	kadar kreatinin
6	urea_darah	kadar urea
7	hemoglobin	kadar Hb
8	kalium	kadar K
9	natrium	kadar Na
10	albumin_serum	kadar albumin
11	protein_urin	level protein urin
12	sel_darah_merah	normal/abnormal
13	penyakit_ginjal_kronis	label (0 = sehat, 1 = CKD)

4.3. *Data Preparation*

Dataset penelitian awalnya terdiri dari 19 atribut, namun setelah proses pembersihan awal, hanya 13 atribut yang digunakan dalam pemodelan. Atribut yang dihapus meliputi id, sel_nanah, hipertensi, diabetes, penyakit_jantung, dan perokok karena tingkat ketidaklengkapan tinggi dan ketidakkonsistenan pencatatan. Dataset yang diberikan juga telah dinormalisasi formatnya serta dinon-identifikasi sehingga tidak mengandung informasi pribadi pasien. Seluruh dataset (500 *record*) digunakan tanpa pembagian manual data latih dan data uji. Evaluasi dilakukan melalui operator *Performance* di *RapidMiner* menggunakan metode resubstitution evaluation, yaitu pengujian model menggunakan data yang sama dengan data pelatihan. Karena itu, *confusion matrix* yang ditampilkan merupakan hasil prediksi model terhadap keseluruhan dataset. Berikut data set yang digunakan dapat dilihat pada Tabel 2.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

Tabel 2. Data Set Penyakit Ginjal Kronis

usia	jenis_kelamin	tekanan_darah	glukosa	kreatinin_serum	urea_darah	hemoglobin	kalium	natrium	albumin	protein_urin	sel_darah_merah	CKD
35	M	135.1	187.1	0.4	35.3	12.7	4.8	136.2	4.66	negative	normal	0
41	M	149.7	97.9	0.4	15.9	14.4	3.45	137.4	3.59	+++	abnormal	0
30	F	163.3	142.8	2.06	26.9	17.4	4.31	136.5	4.67	trace	normal	0
41	F	150.3	185.1	2.74	48.7	15.6	4.76	132.6	3.75	negative	abnormal	1
45	M	93.2	104.8	1.11	37.1	14.5	4.21	135.1	2.69	trace	abnormal	0
48	F	104.4	111.9	1.05	46.7	14.7	5.92	140.1	4.34	negative	normal	0
49	M	117.5	96.7	1.26	74.5	13.6	3.2	139.3	3.45	+	normal	1
38	M	130.5	79.4	2.25	5	14.9	4.84	138.7	4.22	+	normal	1
48	M	140.4	94	0.4	67.3	7.2	4.1	134	3.82	+++	abnormal	1
30	M	115.5	122.4	1.18	37.1	13.3	4.66	137.3	2.87	negative	normal	0
50	M	133.7	121.4	2.08	17	14.1	4.46	129.3	3.9	trace	normal	1
36	M	114.9	89.2	1.65	8.4	10.6	3.76	138.7	5	+	normal	0
47	M	117.8	129.4	1.37	31.2	14.5	4.34	143.7	4.1	negative	abnormal	0
42	F	101.9	60	2.1	53.4	13	3.09	132.1	3.77	trace	normal	1
44	F	111.5	133.2	0.4	32.4	17.9	4.56	135.7	4.45	trace	abnormal	0
49	F	103	153.3	1.81	65.2	17.4	4.38	139.3	3.93	negative	normal	1
43	F	110.5	60	0.61	6	12.3	4.58	143	3.93	negative	normal	0
36	M	151.1	135	0.4	34.8	13.1	4.83	136.2	3.15	++	normal	0
41	M	111	169.1	1.66	10	13.3	4.42	135.2	3.94	++	normal	1
...												

4.4. *Modeling*

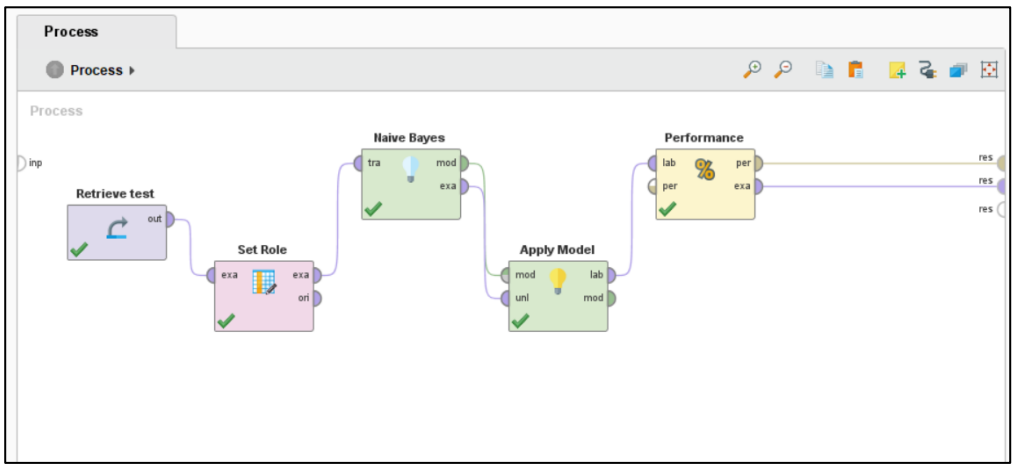
Model dibangun menggunakan operator Naive Bayes untuk menghasilkan struktur klasifikasi yang sesuai dengan karakteristik dataset yang digunakan. Pemilihan algoritma ini dilakukan karena metode Naive Bayes mampu bekerja secara efisien pada data berdimensi tinggi dan memberikan hasil prediksi yang stabil meskipun distribusi datanya sederhana. Proses pemodelan dilakukan dengan mengatur alur operasi pada perangkat *RapidMiner* sehingga setiap tahap pengolahan data dapat dijalankan secara terstruktur. Model dibangun menggunakan operator Naive Bayes dengan struktur proses dapat dilihat pada Gambar 2.

Gambar tersebut menunjukkan alur proses pengolahan data menggunakan algoritma *Naive Bayes* di *RapidMiner*. Proses diawali dengan operator *Retrieve Test*, yang berfungsi mengambil dataset dari *repository* untuk digunakan sebagai data uji. Selanjutnya, data tersebut diproses oleh operator *Set Role*, di mana atribut target (label) ditentukan perannya sebagai "label", sehingga algoritma klasifikasi dapat mengenali atribut mana yang harus diprediksi. Setelah peran atribut disesuaikan, dataset dikirim ke operator *Naive Bayes*, yang bertugas membangun model klasifikasi berdasarkan pola dan probabilitas dari data pelatihan yang tersedia.

Model yang dihasilkan kemudian diterapkan pada data uji melalui operator *Apply Model*, sehingga setiap *record* pada data uji memperoleh prediksi kelas. Hasil prediksi tersebut selanjutnya dievaluasi melalui operator *Performance*, yang menghitung metrik performa seperti akurasi, presisi, *recall*, dan *confusion matrix*. Secara keseluruhan, gambar ini menggambarkan alur lengkap mulai dari pengambilan data, penentuan label, pembangunan model, penerapan model, hingga evaluasi performa untuk menilai kualitas

p-ISSN : 2803-1507 e-ISSN : 2803-1531

model *Naive Bayes* yang dibangun. Evaluasi ini membantu menentukan apakah model sudah layak digunakan atau perlu dilakukan perbaikan lebih lanjut.



Gambar 2. Proses Rapid Miner Naive Bayes

4.5. *Evaluation*

Model dievaluasi menggunakan beberapa metrik performa guna memastikan bahwa hasil prediksi dapat menggambarkan kemampuan algoritma dalam mengenali pola pada data uji. Proses evaluasi dilakukan setelah model selesai dibangun dan diterapkan pada seluruh record, sehingga setiap prediksi dapat dianalisis berdasarkan keluaran numerik maupun kategorikal. Tahapan ini memberikan gambaran mengenai seberapa baik model dalam mengklasifikasikan kelas CKD dan non-CKD, terutama ketika distribusi kelas memiliki variasi tertentu. Hasil proses evaluasi tersebut ditampilkan secara lengkap seperti terlihat pada Gambar 3.

accuracy: 94.40%

	true 0	true 1	class precision
pred. 0	313	21	93.71%
pred. 1	7	159	95.78%
class recall	97.81%	88.33%	

Gambar 3. Hasil Proses Model Naïve Bayes

Secara keseluruhan, model *Naïve Bayes* menunjukkan performa yang tinggi dengan akurasi sebesar 94.40%. Meskipun demikian, terdapat perbedaan yang cukup mencolok antara *recall* untuk kelas CKD dan non-CKD. Nilai *recall* untuk kelas CKD hanya mencapai 88.33%, lebih rendah dibandingkan dengan *recall* kelas non-CKD sebesar 97.81%. Perbedaan ini dapat dijelaskan melalui dua faktor utama.

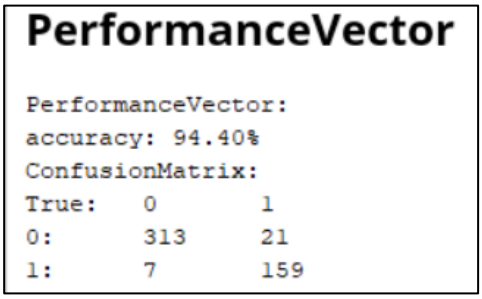
Pertama, jumlah kasus CKD dalam dataset sedikit lebih rendah dibandingkan kasus non-CKD, sehingga distribusi kelas tidak sepenuhnya seimbang. Kondisi ini membuat model lebih mudah mempelajari pola pada kelas non-CKD. Kedua, beberapa atribut klinis menunjukkan nilai yang tumpang tindih (*overlap*) antara pasien CKD dan non-CKD, terutama pada kasus *borderline*. Misalnya, rentang nilai kreatinin serum, albumin, dan urea darah tidak sepenuhnya terpisah antar kelas, sehingga sebagian pasien CKD dengan nilai masih mendekati normal lebih sulit dikenali oleh model.

Berdasarkan tinjauan literatur dan hasil analisis distribusi dataset penelitian ini, terdapat beberapa atribut klinis yang memiliki pengaruh kuat terhadap prediksi CKD. Atribut dengan kontribusi paling besar antara lain kreatinin serum, urea darah, albumin serum, hemoglobin, dan tekanan darah, yang secara medis merupakan indikator utama fungsi ginjal dan kondisi metabolik tubuh. Pola yang ditemukan dalam dataset Jayapura juga konsisten dengan bukti klinis, di mana peningkatan kreatinin dan urea darah serta penurunan albumin dan hemoglobin merupakan penanda kuat terjadinya gangguan fungsi ginjal. Hubungan antara pola perubahan nilai tersebut dengan kondisi fungsi ginjal telah dibuktikan melalui berbagai penelitian klinis yang mengkaji mekanisme progresivitas penyakit.

Selain itu, model dengan akurasi 94.40% yang diperoleh dalam penelitian ini menunjukkan performa yang sejalan dengan berbagai studi terdahulu. Penelitian sebelumnya melaporkan akurasi *Naïve Bayes* pada prediksi CKD berada pada rentang 90% hingga 99%, baik pada dataset UCI maupun dataset klinis. Dengan

p-ISSN : 2803-1507 e-ISSN : 2803-1531

demikian, hasil penelitian ini menguatkan temuan bahwa *Naïve Bayes* merupakan algoritma yang stabil, efisien, dan efektif untuk diagnosis awal CKD, terutama pada dataset dengan kombinasi fitur numerik dan kategorikal. Meskipun penelitian ini tidak membandingkan dengan model lain secara langsung, hasil yang dicapai dapat dianggap valid sebagai *baseline* dan selaras dengan performa metode serupa dalam riset-riset sebelumnya. Pendekatan ini memberikan dasar analitis yang kuat dalam menilai efektivitas model berbasis probabilistic learning pada kasus medis serupa. Hasil tersebut dapat diamati melalui visualisasi evaluasi model seperti ditunjukkan pada Gambar 4.



Gambar 4. Description

Gambar 4 menunjukkan hasil evaluasi model *machine learning* yang ditampilkan dalam bentuk *Performance*Vector. Nilai utama yang ditampilkan adalah accuracy sebesar 94.40%, yang berarti model mampu membuat prediksi yang benar sebanyak 94,40% dari seluruh data uji. Di bawahnya terdapat *Confusion Matrix*, yaitu tabel yang menggambarkan jumlah prediksi benar dan salah untuk masing-masing kelas. Pada baris "0", terlihat bahwa model memprediksi kelas 0 dengan benar sebanyak 313 kali, namun salah memprediksi kelas 0 sebagai kelas 1 sebanyak 21 kali. Sementara itu, pada baris "1", model berhasil memprediksi kelas 1 dengan benar sebanyak 159 kali, dan salah memprediksi kelas 1 sebagai kelas 0 sebanyak 7 kali. Secara keseluruhan, matriks ini memberikan gambaran jelas bahwa model memiliki performa yang sangat baik, dengan jumlah prediksi benar yang dominan pada kedua kelas, serta kesalahan prediksi yang relatif kecil.

4.6. *Deployment*

Model *Naive Bayes* yang dihasilkan memiliki beberapa potensi penerapan yang sangat relevan dalam konteks layanan kesehatan. Dengan performanya yang tinggi dan proses komputasi yang ringan, model ini dapat digunakan sebagai bagian dari sistem pendukung keputusan klinis untuk memberikan prediksi cepat terhadap kemungkinan terjadinya penyakit ginjal kronis (CKD). Kemampuan prediksi ini juga menjadikannya efektif sebagai alat *screening* awal bagi pasien sebelum menjalani pemeriksaan diagnostik yang lebih mendalam, sehingga tenaga medis dapat mengidentifikasi pasien berisiko lebih awal. Selain itu, model ini berpotensi untuk diintegrasikan ke dalam aplikasi monitoring kesehatan, memungkinkan pengguna atau tenaga kesehatan melakukan pemantauan kondisi ginjal secara berkala. Dalam skenario yang lebih lanjut, model dapat digunakan untuk prediksi secara *real-time* berdasarkan parameter klinis yang diinput oleh pasien atau perangkat medis. Karena algoritma *Naive Bayes* bekerja dengan kompleksitas komputasi yang rendah, implementasi model ini sangat cocok untuk berbagai platform, mulai dari aplikasi desktop hingga aplikasi berbasis *web*, tanpa membutuhkan sumber daya pemrosesan yang besar.

5. DISKUSI

Temuan penelitian menunjukkan bahwa metode demonstrasi mampu meningkatkan kemampuan praktik wudhu anak secara konsisten, sejalan dengan penjelasan dalam [3] mengenai efektivitas pembelajaran berbasis pengamatan langsung. Peningkatan skor pada setiap pertemuan menunjukkan bahwa stimulus visual dan praktik terstruktur memberikan pemahaman yang lebih konkret bagi anak usia dini. Pola respons ini juga memperkuat laporan dalam [7] yang menyoroti peran aktivitas *hands-on* dalam mempercepat perkembangan keterampilan motorik halus. Hasil tersebut menunjukkan bahwa penyajian langkah-langkah wudhu secara terarah memberikan kejelasan prosedural yang mudah dipahami anak.

Perbandingan dengan penelitian sebelumnya menunjukkan bahwa metode demonstrasi memiliki stabilitas hasil yang serupa dengan temuan dalam [5], khususnya dalam konteks pembelajaran ibadah yang membutuhkan urutan sistematis. Meskipun demikian, terdapat beberapa anak yang membutuhkan waktu lebih lama untuk meniru gerakan wudhu, suatu kondisi yang berbeda dari gambaran peningkatan cepat yang

p-ISSN : 2803-1507 e-ISSN : 2803-1531

dilaporkan dalam [8]. Variasi ini mengindikasikan bahwa kesiapan perkembangan anak menjadi faktor penting dalam efektivitas metode demonstrasi. Temuan tersebut juga menunjukkan bahwa karakteristik individu dapat menghasilkan kecepatan perkembangan yang tidak selalu seragam.

Beberapa fenomena menarik muncul pada tahap awal pembelajaran, terutama ketika anak menunjukkan kebingungan terhadap urutan gerakan, selaras dengan catatan dalam [10] mengenai tantangan memori prosedural pada usia dini. Ketidakkonsistenan ini memperlihatkan bahwa proses internalisasi langkah wudhu membutuhkan pengulangan yang lebih intensif. Selain itu, ketertarikan anak terhadap aktivitas air memperkuat temuan dalam [12] mengenai pengaruh stimulus sensorik terhadap peningkatan fokus belajar. Respons tersebut menunjukkan bahwa elemen sensorik berperan dalam memperkuat pemahaman anak terhadap aktivitas ritual.

Implikasi teoretis penelitian ini memperluas pemahaman tentang bagaimana metode demonstrasi mendukung pembelajaran prosedural melalui integrasi pengamatan dan praktik. Dari sisi praktis, metode ini dapat digunakan sebagai alternatif pembelajaran ibadah yang efektif bagi guru PAUD dalam memperkenalkan tahapan wudhu secara sistematis. Penelitian ini juga memberikan gambaran bahwa pembelajaran ritual melalui demonstrasi mampu menciptakan suasana pembelajaran yang lebih aktif dan mudah diikuti oleh anak. Meskipun begitu, penelitian ini memiliki keterbatasan pada jumlah subjek dan variasi konteks pembelajaran sehingga interpretasi hasil perlu disesuaikan dengan cakupan penelitian yang terbatas.

6. KESIMPULAN DAN SARAN

6.1. *Kesimpulan*

Penelitian ini berhasil membangun model prediksi penyakit ginjal kronis (CKD) menggunakan algoritma *Naïve Bayes* dengan memanfaatkan dataset klinis yang telah melalui proses pembersihan sebelumnya. Model menunjukkan performa yang sangat baik, dengan akurasi sebesar 94.40%, *precision* CKD sebesar 95.78%, dan *recall* CKD sebesar 88.33%. Hal ini membuktikan bahwa *Naïve Bayes* mampu mengklasifikasikan pasien CKD maupun non-CKD dengan tingkat ketepatan yang tinggi, terutama dalam mengenali pasien sehat, meskipun masih memiliki keterbatasan dalam mendeteksi beberapa kasus CKD pada kondisi borderline. Temuan ini memberikan gambaran mengenai perilaku model ketika berhadapan dengan variasi karakteristik pasien pada kondisi klinis nyata.

Kontribusi utama penelitian ini terletak pada pemanfaatan data klinis nyata dari Kota Jayapura, sehingga model yang dikembangkan benar-benar mencerminkan karakteristik kesehatan masyarakat Papua. Sebagian besar studi sebelumnya menggunakan dataset dari luar Papua, sehingga penelitian ini memberikan nilai tambah yang signifikan bagi literatur dan praktik kesehatan lokal. Hasil penelitian ini juga menjadi landasan awal bagi pengembangan sistem deteksi dini CKD yang lebih relevan dan adaptif bagi layanan kesehatan di Jayapura serta wilayah Indonesia Timur. Aspek ini menunjukkan bahwa pemanfaatan data lokal berpotensi meningkatkan kecocokan model terhadap kebutuhan pelayanan kesehatan berbasis wilayah.

6.2. *Saran*

Untuk meningkatkan generalisasi dan keandalan model, penelitian selanjutnya disarankan menggunakan dataset yang lebih besar dan lebih bervariasi, khususnya untuk menangani kasus CKD borderline. Selain itu, validasi klinis langsung perlu dilakukan melalui kolaborasi dengan tenaga medis di fasilitas kesehatan Jayapura, sehingga model dapat diuji dengan data real-time dan dievaluasi berdasarkan diagnosis aktual. Pendekatan ini penting agar model benar-benar siap diterapkan dalam konteks klinis di lapangan. Pengembangan lanjutan juga dapat mengeksplorasi algoritma *machine learning* lain atau pendekatan *ensemble* guna meningkatkan performa di atas *Naïve Bayes*. Integrasi model ke dalam aplikasi klinis maupun sistem informasi kesehatan daerah juga direkomendasikan, sehingga dapat dimanfaatkan sebagai alat screening awal yang cepat, ringan, dan efektif untuk membantu pengambilan keputusan medis dan pemantauan kondisi pasien.

DAFTAR PUSTAKA

[1] M. Rizal, M. Z. Syahaf, S. R. Priyambodo, and Y. Ramdhani, “Optimasi Algoritma Naïve Bayes Menggunakan Forward Selection untuk Klasifikasi Penyakit Ginjal Kronis,” *Naratif J. Nas. Riset, Apl. dan Tek. Inform.*, vol. 5, no. 1, pp. 71–80, 2023, doi: 10.53580/naratif.v5i1.200.

[2] V. Wulandari, W. J. Sari, and Z. Alfian, “Implementasi Algoritma Naïve Bayes Classifier dan K-

Nearest Neighbor untuk Klasifikasi Penyakit Ginjal Kronik," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 2, pp. 710–718, 2024, doi: 10.57152/malcom.v4i2.1229.

[3] Taryadi, E. Yunianto, and Kasmari, "Diagnostik Penyakit Ginjal Kronis Menggunakan Model Klasifikasi Support Vector Machine," *IC Tech Maj. Ilm.*, vol. 19, no. 1, pp. 39–44, 2024, doi: 10.47775/ictech.v19i1.291.

[4] A. K. Wijaya, R. Syifa, I. N. Rahmadianto, and R. K. Hapsari, "Identifikasi Penyakit Ginjal Kronis Menggunakan Algoritma K-Nearest Neighbour (k-NN)," in *Prosiding Seminar Nasional Informatika Bela Negara*, 2024, pp. 361–365. [Online]. Available: https://santika.upnjatim.ac.id/submissions/index.php/santika/article/view/504

[5] N. Surojudin, "Perbandingan Metode Klasifikasi dalam Memprediksi Penyakit Ginjal Kronis," *J. Inform. Ekon. Bisnis*, vol. 7, no. 3, pp. 715–723, 2025, doi: 10.37034/infeb.v7i3.1263.

[6] I. Ikko, M. Rizky, S. Yusuf, and I. Sriyanto, "Perbandingan Kinerja Algoritma Naive Bayes, Support Vector Machine dan Random forest untuk Prediksi Penyakit Ginjal Kronis," *Pros. Semin. Nas. Darmajaya*, vol. 18, no. 0, pp. 139–151, 2023, [Online]. Available: https://jurnal.darmajaya.ac.id/index.php/PSND/article/view/3832

[7] Irwan Reza Pahlevi, "Penerapan Naive Bayes Untuk Prediksi Penyakit Diabetes dengan Menggunakan Rapid Miner," *J-CEKI J. Cendekia Ilm.*, vol. 4, no. 4, pp. 313–322, 2025, doi: 10.56799/jceki.v4i4.8531.

[8] A. Arman, P. Prasetya, F. N. Arifany, F. B. Pradnyaparamita, and J. Laksito, "A Digital Printing Application as an Expression Identification System," *J. Technol. Informatics Eng.*, vol. 1, no. 2, pp. 5–15, Jul. 2022, doi: 10.51903/jtie.v1i2.135.

[9] D. Adi, S. Rahardjo, M. Andriana, and S. H. Prasastono, "Consumer Behavior and Purchase Intentions: A Machine Learning-Based Sentiment Analysis of Social Media Data," *J. Manag. Informatics*, vol. 3, no. 1, pp. 157–173, Apr. 2024, doi: 10.51903/jmi.v3i1.45.

[10] Y. Raudha, M. G. Rohman, and Munif, "Sistem Pakar Diagnosa Penyakit pada Sapi Menggunakan Naive Bayes," *J. Ilm. Sist. Inf.*, vol. 4, no. 2, pp. 259–275, Nov. 2025, doi: 10.51903/pmxpdm56.

[11] A. Qurotul, E. Tasia, N. Nazira, P. F. Pratama, M. R. Anugrah, and J. Adhiva, "Implementasi Algoritma Naïve Bayes Classifier (NBC) untuk Klasifikasi Penyakit Ginjal Kronik," *J. Sist. Komput. dan Inform.*, vol. 4, no. 1, pp. 72–76, 2022, doi: 10.30865/json.v4i1.4781.

[12] R. Tyasnurita and S. W. Hapsari, "Identification of Chronic Kidney Disease Using Naive Bayes, Adaboost, and Random Forest Learning Methods," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 6, no. 1, pp. 115–120, 2020, doi: 10.33480/jitk.v6i1.1403.

[13] N. S. Abd and D. A. Abdullah, "Diagnose of Chronic Kidney Diseases by Using Naive Bayes Algorithm," *J. Al-Qadisiyah Comput. Sci. Math.*, vol. 13, no. 2, pp. 46–55, 2021, doi: 10.29304/jqcm.2021.13.2.819.

[14] K. N. Qodri, M. R. Fikri, and L. Ardi, "Prediksi Analitik untuk Penyakit Ginjal Kronis: Perbandingan Metode Machine Learning," *J. Keilmuan Teknol. Inf.*, vol. 1, no. 1, pp. 15–22, 2025, doi: 10.61902/jkti.v1i1.1686.

[15] N. P. Nugraha, R. Azim, S. Z. Daffa, and P. S. Ningayu, "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, doi: 10.36706/jres.v5i1.63.

[16] M. Dhanush, H. Raj, P. Bothra, and R. Zanwar, "Chronic Kidney Disease Prediction by using Naive Bayes," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 9, no. 5, pp. 386–390, 2024, doi: 10.22214/ijraset.2024.62002.

[17] A. R. Richardo Anggara, "Implementasi Algoritma DBSCAN Dalam Mengelompokan Data Pasien Terdiagnosa Penyakit Ginjal Kronis (PGK)," *J. Algoritm.*, vol. 3, no. 1, pp. 114–122, 2022, doi: 10.35957/algoritme.v3i1.3593.

p-ISSN : 2803-1507 e-ISSN : 2803-1531

[18] A. R. Wibowo, Y. P. Astuti, E. Kartikadarma, and E. R. Subhiyakto, "Penerapan Metode Naïve Bayes Classifier untuk Klasifikasi Sentimen pada Judul Berita," *LogicLink J. Artif. Intell. Multimed. Informatics*, vol. 1, no. 1, pp. 1–12, 2024, doi: 10.28918/logiclink.v1i1.7684.

[19] R. M. Sari, V. Tasril, S. Wahyuni, and S. E. Putri, "Desain Aplikasi Naïve Bayes untuk Memprediksi Kelulusan Siswa," *Pros. Semin. Nas. Teknol. Inov. dan Kolaborasi Disiplin Ilmu*, vol. 1, no. 1, pp. 4–9, 2024, [Online]. Available: https://journal.fkpt.org/index.php/sinekad/issue/current

[20] I. Gaho and A. Maslan, "Implementasi Data Mining untuk Memprediksi Penjualan Produk Terlaris pada Petshop Menggunakan Algoritma Naive Bayes," *Comput. Sci. Ind. Eng.*, vol. 11, no. 2, pp. 31–40, 2024, doi: 10.33884/comasiejournal.v11i2.9041.

[21] T. Tupari, S. Abdullah, and C. Chairani, "Visualisasi Data Analisa Sentimen RUU Omnibus Law Kesehatan Menggunakan KNN dengan Software RapidMiner," *J. Inform. J. Pengemb. IT*, vol. 8, no. 3, pp. 261–268, 2023, doi: 10.30591/jpit.v8i3.5641.

[22] W. M. Putri, E. Budianita, F. Syafria, and I. Afrianty, "Klasifikasi Penyakit Ginjal Kronis Menggunakan Information Gain dan LVQ," *J. Inf. Syst. Manag.*, vol. 7, no. 1, pp. 48–56, 2025, doi: 10.24076/joism.2025v7i1.2102.

[23] P. A. R. Devi, "Klasifikasi Penyakit Gagal Ginjal Kronis dengan Metode KNN (Studi Kasus RS di Kab Gresik)," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 9, no. 3, pp. 1739–1748, 2024, doi: 10.29100/jipi.v9i3.6226.

[24] I. I. M. Rizky, W. F. Purwaningtyas, and M. Rahmawati, "Analisis Komparatif Algoritma Klasifikasi untuk Prediksi Dini Penyakit Ginjal Kronis," *J. Kesehat. dan Sist. Inf.*, vol. 9, no. 2, pp. 89–95, 2025, doi: 10.55919/jk.v9i2.207.

[25] S. N. Chotimah and A. R. Rozzaqi, "Klasifikasi Diagnosis Penyakit Ginjal Kronis Dengan Menerapkan Konsep Algoritma Naãve Bayes," *JIPETIKJurnal Ilm. Penelit. Teknol. Inf. Komput.*, vol. 4, no. 1, pp. 8–15, 2023, doi: 10.26877/jipetik.v4i1.16174.

[26] T. Arifin and D. Ariesta, "Prediksi Penyakit Ginjal Kronis Menggunakan Algoritma Naive Bayes Classifier Berbasis Particle Swarm Optimization," *J. Tekno Insentif*, vol. 13, no. 1, pp. 26–30, 2019, doi: 10.36787/jti.v13i1.97.

[27] Y. N. M. Hendri Hartono, Alyauma Hajjah, "Application of the Naïve Bayes Classifier Method for News Title Classification," *J. Simantec*, vol. 12, no. 1, pp. 37–46, 2023, [Online]. Available: https://journal.trunojoyo.ac.id/simantec/article/viewFile/19398/9506

[28] H. Mukaromah, "Perbandingan Kinerja Algoritma Naïve Bayes dan Decision Tree untuk Prediksi Penyakit Batu Ginjal," *J. Rekayasa Perangkat Lunak*, vol. 4, no. 1, pp. 9–14, 2025, [Online]. Available: https://journal.aisyahuniversity.ac.id/index.php/J-Rapa/article/download/1981/923/4406