

ANALISIS POLA PENYAKIT BERDASARKAN DATA PASIEN DI UPT PUSKESMAS TANJUNG TIRAM MENGGUNAKAN ALGORITMA APRIORI

Riski Perdamaian Ndraha^{1*}, Dicky Apdillah¹

Universitas Asahan, Asahan

e-mail: ¹riskidamai4@gmail.com, ¹dicky@nusa.net.id

Abstract: *This study discusses the application of the Apriori algorithm to analyze the pattern of relationships between diseases in patient medical record data at UPT Puskesmas Tanjung Tiram. Medical records serve as an important source of information to support medical decision-making, particularly in identifying diseases that frequently occur together. This study utilized 365 outpatient medical record data, which were processed using Python and integrated with PHP as a web-based interface. The parameters used in the Apriori algorithm include a minimum support of 10%, a minimum confidence of 50%, and a minimum lift of 1.2. The results of the analysis revealed four significant disease patterns, namely hypertension → fever, acute respiratory infection (ARI) → fever, gout → fever, and cholesterol → fever, indicating strong associations among these diseases. The developed system can display analytical results quickly, accurately, and interactively, making it useful in supporting diagnosis processes and medical decision-making at healthcare centers. This research is expected to contribute to the utilization of data mining for disease pattern analysis and the development of web-based health recommendation systems.*

Keyword: *apriori algorithm, data mining, PHP, python, disease patterns, medical records*

Abstrak: Penelitian ini membahas penerapan algoritma Apriori untuk menganalisis pola keterkaitan antar penyakit pada data rekam medis pasien di UPT Puskesmas Tanjung Tiram. Data rekam medis merupakan sumber informasi penting dalam mendukung pengambilan keputusan medis, terutama untuk mengidentifikasi penyakit yang sering muncul secara bersamaan. Penelitian ini menggunakan 365 data rekam medis pasien rawat jalan yang diolah dengan bantuan Python dan diintegrasikan dengan PHP sebagai antarmuka web. Parameter yang digunakan pada algoritma Apriori meliputi *minimum support* sebesar 10%, *minimum confidence* sebesar 50%, dan *minimum lift* sebesar 1.2. Hasil analisis menunjukkan terdapat empat pola penyakit signifikan, yaitu hipertensi → demam, ISPA → demam, asam urat → demam, dan kolesterol → demam, dengan tingkat keterkaitan yang kuat. Sistem yang dibangun mampu menampilkan hasil analisis secara cepat, akurat, dan interaktif, sehingga dapat dimanfaatkan untuk mendukung proses diagnosis dan pengambilan keputusan medis di lingkungan Puskesmas. Penelitian ini diharapkan memberikan kontribusi dalam pemanfaatan data mining untuk analisis pola penyakit serta pengembangan sistem rekomendasi kesehatan berbasis web.

Kata kunci: algoritma apriori, data mining, PHP, python, pola penyakit, rekam medis

PENDAHULUAN

Dalam Peraturan Menteri Kesehatan Republik Indonesia Nomor 75 Tahun 2014, Puskesmas adalah fasilitas pelayanan kesehatan tingkat pertama yang

bertugas melaksanakan upaya kesehatan masyarakat dan pelayanan kesehatan individu (Permenkes, 2014). Puskesmas lebih mengutamakan upaya pencegahan penyakit dan penyuluhan kesehatan untuk

meningkatkan kesehatan masyarakat secara maksimal di wilayah tugasnya.

Namun demikian, di UPT Puskesmas Tanjung Tiram masih terdapat tantangan serius terkait belum tersedianya sistem komputarisasi pada pengelolaan rekam medis. Selama ini, pencatatan data pasien dilakukan secara manual dan hanya sebatas fungsi administratif seperti pencatatan identitas dan riwayat kunjungan. Kondisi ini membuat data kesehatan yang sebenarnya sangat berharga tidak dapat dimanfaatkan untuk mendeteksi pola penyakit, menemukan hubungan antarpenyakit, atau merancang strategi pencegahan yang tepat. Seiring meningkatnya jumlah pasien setiap tahun, permasalahan ini semakin mendesak karena data yang terkumpul tidak terolah secara optimal, sehingga pelayanan lebih berfokus pada aspek kuratif dan belum didukung bukti berbasis data untuk memperkuat upaya preventif dalam meningkatkan kesehatan masyarakat.

Selain itu, proses diagnosis penyakit dalam praktiknya juga kerap menghadapi kesulitan karena gejala penyakit sering kali saling tumpang tindih dan berhubungan dengan berbagai macam penyakit (Rusmina et al., 2024). Kompleksitas ini menuntut adanya pendekatan yang sistematis dan berbasis data untuk memetakan keterkaitan antar gejala dan penyakit. (Ma et al., 2022) menyatakan bahwa satu gejala dapat terkait dengan lebih dari satu jenis penyakit, sehingga diperlukan pengolahan data rekam medis secara terstruktur untuk membantu proses identifikasi penyakit. (Kristanto & Indonesia, 2023) menegaskan bahwa metode analisis data dapat mendukung diagnosis dan perencanaan penanganan medis secara lebih akurat, khususnya di fasilitas layanan kesehatan dasar seperti Puskesmas.

Salah satu pendekatan yang dapat digunakan untuk mengolah data rekam medis adalah data mining. Data mining merupakan proses komputasi yang bertujuan menemukan pola dan informasi bermakna dari data berukuran besar

(Ritha et al., 2021). (Leman, 2024) menjelaskan bahwa data mining memanfaatkan teknik statistik, kecerdasan buatan, matematika, dan pembelajaran mesin untuk mengidentifikasi pengetahuan tersembunyi dari kumpulan data. Semakin besar volume data yang tersedia, semakin besar pula potensi informasi yang dapat dieksplorasi. Dalam konteks kesehatan, data mining dapat digunakan untuk mengidentifikasi faktor yang mempengaruhi diagnosis penyakit serta menggali pola kemunculan penyakit (Nuryawan, 2024).

Beberapa penelitian sebelumnya menunjukkan efektivitas data mining dalam lingkungan pelayanan kesehatan. Penelitian oleh (Aslyza, 2025) membuktikan bahwa algoritma Apriori dapat digunakan untuk menemukan hubungan antarpenyakit yang sering muncul bersamaan, sehingga rumah sakit dapat melakukan langkah pencegahan, meningkatkan ketersediaan obat, serta perencanaan fasilitas medis dengan lebih baik. Temuan serupa juga dijelaskan oleh (Luhur et al., 2020) yang menunjukkan bahwa algoritma Apriori efektif dalam mengungkap pola keterkaitan gejala penyakit dalam dataset medis, sehingga membantu diagnosis dini dan pengelolaan penyakit.

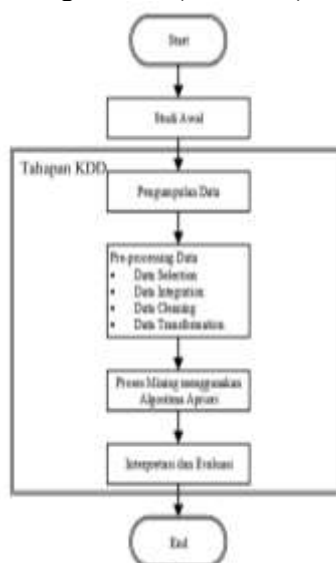
Algoritma Apriori merupakan salah satu teknik data mining yang banyak digunakan untuk menemukan pola asosiasi atau hubungan antaritem dalam jumlah data besar (Rusmina et al., 2024). Algoritma ini bekerja dengan menghitung frekuensi kemunculan data (*support*) dan menghitung kemungkinan keterhubungan antar item (*confidence*) untuk menghasilkan pola atau aturan asosiasi (Ananda et al., 2023). Kelebihan algoritma Apriori antara lain struktur yang sederhana, mudah dipahami, bersifat unsupervised, serta mampu menghasilkan seluruh aturan asosiasi yang memenuhi batas support dan confidence tertentu (Kristanto & Indonesia, 2023). Dalam konteks kesehatan, algoritma ini dapat digunakan untuk menganalisis keterkaitan antar gejala dan penyakit yang sering

muncul bersamaan sehingga mendukung proses diagnosis dan pengambilan keputusan medis secara lebih cepat dan akurat.

Dengan permasalahan yang dihadapi UPT Puskesmas Tanjung Tiram dan dukungan berbagai penelitian sebelumnya, algoritma Apriori dipandang potensial untuk membantu mengolah rekam medis menjadi informasi pola penyakit yang dapat digunakan dalam perencanaan program kesehatan, deteksi dini potensi wabah, serta penguatan edukasi kesehatan yang lebih tepat sasaran.

METODE

Penelitian ini dilaksanakan melalui serangkaian tahapan sistematis yang diawali dengan studi pendahuluan, kemudian dilanjutkan dengan proses *Knowledge Discovery in Databases* (KDD), yang meliputi pengumpulan data, pra-proses data, penambangan data menggunakan algoritma Apriori, serta tahapan interpretasi dan evaluasi hasil. Seluruh tahapan tersebut disusun dalam suatu kerangka kerja penelitian yang berfungsi sebagai pedoman konseptual dalam pelaksanaan penelitian. Adapun kerangka kerja penelitian divisualisasikan melalui diagram alir (flowchart) berikut:



Gambar 1 Tahapan Penelitian

Tahapan penelitian ini disusun untuk memberikan alur yang jelas dan terstruktur dalam pelaksanaan penelitian. Tahapan penelitian diawali dengan studi pendahuluan, yang bertujuan untuk mengidentifikasi permasalahan, menentukan ruang lingkup penelitian, dan menelaah literatur terkait metode *data mining*, khususnya algoritma Apriori. Tahapan awal ini menjadi dasar dalam merumuskan kebutuhan data dan menetapkan pendekatan analisis yang akan digunakan.

Selanjutnya, penelitian memasuki proses *Knowledge Discovery in Databases* (KDD) yang terdiri dari beberapa langkah utama. Tahap pertama adalah pengumpulan data dari meliputi metode observasi, wawancara dan studi literatur, untuk mengumpulkan data rekam medis yang relevan dari UPT Puskesmas Tanjung Tiram. Setelah data terkumpul, dilakukan pra-processing data, yang meliputi seleksi data, pembersihan data, integrasi data, serta transformasi data ke dalam format yang sesuai untuk proses penambangan data.

Tahap berikutnya adalah proses mining, yaitu penerapan algoritma Apriori untuk menemukan pola hubungan antar diagnosis atau variabel yang terdapat dalam dataset. Algoritma Apriori digunakan untuk menghasilkan *frequent itemset* dan *association rules* berdasarkan nilai support dan confidence yang telah ditentukan sebelumnya. Proses ini merupakan inti dari penelitian karena menghasilkan pola-pola asosiasi yang dapat memberikan informasi bermakna.

Selanjutnya, dilakukan interpretasi dan evaluasi hasil, yaitu menganalisis aturan asosiasi yang diperoleh untuk menilai tingkat keterandalan dan relevansinya terhadap konteks data rekam medis. Evaluasi dilakukan dengan mempertimbangkan parameter seperti support, confidence, dan lift, serta kesesuaian temuan dengan kondisi lapangan.

Rangkaian tahapan tersebut disusun dalam bentuk diagram alir (*flowchart*) sebagai representasi visual dari kerangka

kerja penelitian. *Flowchart* ini mempermudah pemahaman mengenai urutan langkah-langkah penelitian serta

keterkaitan antar tahapan dalam keseluruhan proses KDD.

Tabel 1 Format Tabular

Kode Pasien	Hipertensi	Diabetes	ISPA	Anemia	Asam Urat	Gastritis	TB Paru	Kolesterol
P001	1	1	0	0	0	1	0	1
P002	1	0	1	1	0	0	0	0
P003	0	1	1	0	0	1	0	0
P004	1	1	0	0	1	0	0	1
P005	0	0	1	1	0	0	1	0
P006	0	0	0	1	1	0	0	0
P007	1	1	0	0	1	1	0	1
P008	0	0	1	0	0	1	1	0
P009	1	0	0	1	0	0	0	1
P010	0	1	1	0	1	0	0	0

HASIL DAN PEMBAHASAN

Implementasi sistem dilakukan dengan mengintegrasikan algoritma Apriori berbasis Python dengan aplikasi web menggunakan PHP. Python dimanfaatkan untuk melakukan preprocessing, perhitungan frequent itemset, serta pembentukan association rules berdasarkan parameter minimum support 10%, minimum confidence 50%, dan lift di atas 1,2.



Gambar 2 Antarmuka Dashboard

Hasil analisis kemudian diekspor dalam format JSON/CSV agar dapat dibaca oleh sistem PHP. Antarmuka aplikasi dikembangkan untuk mendukung proses unggah dataset rekam medis, pengaturan parameter analisis, pelaksanaan algoritma Apriori, dan penayangan hasil dalam bentuk tabel yang

interaktif. Seluruh komponen antarmuka mulai dari halaman login, dashboard, dataset, hingga analisis berfungsi sebagai sarana untuk mempermudah pengguna mengakses dan memahami hasil pengolahan data.

Hasil implementasi menunjukkan bahwa sistem mampu menampilkan pola penyakit secara akurat dan interaktif. Pola yang paling dominan adalah hubungan antara hipertensi, ISPA, kolesterol, asam urat, dan demam, dengan nilai support, confidence, dan lift yang memenuhi kriteria. Hal ini menunjukkan bahwa algoritma Apriori berhasil mengidentifikasi penyakit yang sering muncul bersamaan.



Gambar 3 Antarmuka Dataset

Pengujian Sistem dilakukan untuk menilai akurasi aturan, performa eksekusi, serta kemudahan penggunaan sistem. (1)

pengujian akurasi aturan, validasi manual terhadap 365 data rekam medis menunjukkan bahwa aturan (*Hipertensi*) ke (*Demam*) merupakan pola yang kuat dengan support 18%, *confidence* 94,1%, dan lift 1,2. Sebaliknya, aturan seperti (*Flu*) ke (*Asam Urat*) tidak memenuhi ambang batas karena memiliki support dan *confidence* yang rendah, sehingga dianggap tidak signifikan. (2) pengujian performa, Proses analisis berjalan sangat cepat dan hasil dapat ditampilkan secara instan melalui antarmuka web. Hal ini menunjukkan bahwa integrasi Python PHP memiliki performa baik untuk data berskala kecil hingga menengah. (3) pengujian keterpakaian (*Usability*), Uji coba oleh pengguna non-teknis menunjukkan bahwa alur penggunaan sistem tergolong mudah: pengguna hanya perlu memasukkan parameter, menjalankan analisis, dan melihat hasil aturan asosiasi. Beberapa masukan disampaikan agar ditambahkan visualisasi grafik untuk mempermudah pemahaman hasil. (4) pengujian fungsionalitas, Seluruh fitur aplikasi, mulai dari unggah data CSV, eksekusi script Python melalui *shell_exec()*, hingga tampilan hasil analisis, bekerja dengan baik tanpa error. Analisis menggunakan parameter dasar menghasilkan empat pola penyakit signifikan yang dapat digunakan sebagai dasar rekomendasi dalam proses diagnosis dan pemantauan kondisi pasien di UPT Puskesmas Tanjung Tiram.



Penyakit	Parameter	Support	Confidence	Lift
Hipertensi	Demam	18%	94,1%	1,2
Hipertensi	ISPA	18%	94,1%	1,2
Demam	Hipertensi	18%	94,1%	1,2
Demam	ISPA	18%	94,1%	1,2

Gambar 4 Hasil Pengujian

Secara keseluruhan, sistem berhasil mengolah data rekam medis menjadi informasi bermakna menggunakan algoritma Apriori. Hasil pola penyakit

yang ditemukan memberikan gambaran kuat mengenai keterkaitan antar penyakit, khususnya hubungan yang melibatkan demam. Implementasi ini terbukti efektif dalam mendukung proses analisis serta pengambilan keputusan medis berbasis data.

SIMPULAN

Rangkaian penelitian yang telah dilakukan, mulai dari pengumpulan data, implementasi algoritma, integrasi sistem, hingga tahap pengujian, diperoleh beberapa simpulan sebagai berikut:

1. Implementasi algoritma Apriori menggunakan Python mampu menghasilkan pola asosiasi penyakit yang akurat dan relevan. Dari 365 data rekam medis rawat jalan, algoritma Apriori berhasil mengidentifikasi hubungan antar penyakit dengan parameter minimum support 10%, minimum confidence 50%, dan minimum lift 1.2. Hasil ini menunjukkan bahwa metode data mining khususnya association rule mining efektif digunakan untuk menganalisis pola kejadian penyakit pada layanan kesehatan tingkat pertama.
2. Integrasi antara Python dan PHP berjalan optimal dan mendukung alur analisis secara otomatis. Proses eksekusi algoritma melalui *shell_exec()*, pertukaran data menggunakan format *JSON/CSV*, serta penyajian hasil melalui antarmuka web terbukti dapat bekerja secara stabil. Integrasi ini memungkinkan pengguna non-teknis untuk melakukan analisis Apriori tanpa harus menjalankan perintah secara manual di Python.
3. Empat pola penyakit signifikan berhasil ditemukan dan memenuhi seluruh parameter analisis. Aturan asosiasi yang lolos seleksi adalah:
Hipertensi → Demam
ISPA → Demam

Asam Urat → Demam
Kolesterol → Demam

Pola Asam Urat → Demam memiliki nilai confidence tertinggi (96,49%), menandakan hubungan yang sangat kuat. Sementara itu, pola ISPA → Demam memiliki support tertinggi (0,227), menunjukkan bahwa pola tersebut paling sering muncul dalam dataset. Nilai lift lebih dari 1 pada seluruh aturan menegaskan bahwa hubungan antar penyakit bukanlah kebetulan semata, tetapi memiliki kaitan yang signifikan. Aplikasi analisis Apriori yang dikembangkan terbukti fungsional, responsif, dan mudah digunakan.

DAFTAR PUSTAKA

- Ananda, Y., Sembiring, B., & Sembiring, E. A. (2023). *Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang*. 1(1), 1–8.
- Aslyza, M. (2025). *MANAJEMEN DATA BERBASIS DATABASE: SOLUSI UNTUK PENYIMPANAN DAN AKSES DATA YANG LEBIH EFISIEN*. 2(3), 909–917.
- Kristanto, K., & Indonesia, U. K. (2023). *Analisis data kesehatan menggunakan big data pada bidang kesehatan*. December.
- Leman, D. (2024). *Sistem Cerdas Rekomendasi Klinik Pratama di Kota Medan Berbasis Data Mining Dengan Metode K-Means Untuk Pasien BPJS dan Umum*. September, 204–214.
- Luhur, U. B., Raya, J. C., Utara, P., & Selatan, J. (2020). *Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Meningkatkan Pola Penjualan Obat*. 1,2. 7(2).
- Ma, H., Ding, J., Liu, M., & Liu, Y. (2022). *Connections between Various Disorders : Combination Pattern Mining Using Apriori Algorithm Based on Diagnosis Information from Electronic Medical Records*. 2022.
- Nuryawan, F. (2024). *Analisis pola minat konsumen dengan algoritma apriori*. 15(2), 269–276.
- Permenkes. (2014). *Permenkes No 75 Tahun 2014*.
- Ritha, N., Suswaini, E., & Pebriadi, W. (2021). *Penerapan Association Rule Menggunakan Algoritma Apriori Pada Poliklinik Penyakit Dalam (Studi Kasus : Rumah Sakit Umum Daerah Bintan)*. 7(November), 222–230.
<https://doi.org/10.34128/jsi.v7i2.329>
- Rusmina, E., Sihombing, V., & Juledi, A. P. (2024). *Analisis Keterkaitan Antara Gejala Penyakit Menggunakan Algoritma Apriori dalam Bidang Kesehatan*. 7, 337