

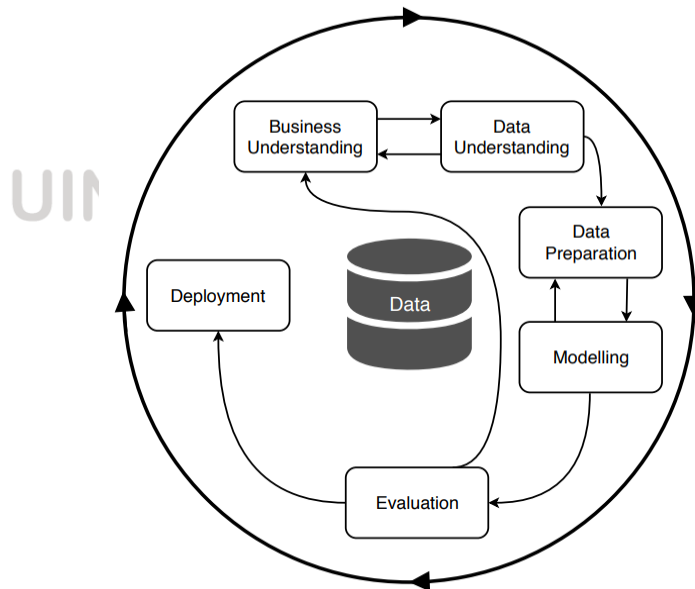
BAB III

METODE PENELITIAN

A. Jenis Penelitian

Penelitian ini menggunakan jenis penelitian kuantitatif dengan pendekatan komparatif. Pendekatan komparatif dipilih karena tujuan penelitian ini adalah membandingkan kinerja dua algoritma *association rule mining*, yaitu algoritma Apriori dan FP-Growth, dalam menemukan pola kombinasi antara item pada data transaksi penjualan.

Penelitian ini menggunakan pendekatan *data mining* dengan kerangka kerja CRISP-DM (*Cross-Industry Standard Process for Data mining*) sebagai metode utama dalam proses pengolahan data. Hasil penelitian berupa data yang telah tervalidasi, informasi kombinasi obat, serta rekomendasi pengelolaan tata letak dan stok obat yang diperoleh dari analisis data transaksi penjualan Apotek Gretta Solok periode Oktober 2025 hingga Februari 2026. Adapun kerangka penelitian dengan CRISP-DM ini akan diuraikan pada gambar 3.1 berikut ini.



Gambar 3. 1 Metode CRISP-DM
Sumber: (Martínez-Plumed dkk., 2021)

Berdasarkan Gambar 3.1 proses CRISP-DM dimulai dari tahap *Business Understanding* dan berlanjut secara berurutan hingga tahap *Deployment*. Setiap tahapan saling berhubungan dan memungkinkan adanya pengulangan (iterasi), sehingga proses *data mining* bersifat siklik (prosesnya berulang) dan berkelanjutan sesuai dengan kebutuhan dan tujuan bisnis. Ada pun penjelasan tiap tahapannya ialah sebagai berikut:

a. *Business Understanding* (Pemahaman Bisnis)

Pada tahap ini dilakukan identifikasi masalah bahwa Apotek Gretta Solok masih menggunakan pencatatan manual dan belum memanfaatkan data transaksi untuk pengambilan keputusan. Selama ini data transaksi hanya disimpan dalam bentuk pencatatan manual sehingga belum dapat dianalisis untuk mengetahui pola pembelian obat.

Berdasarkan kondisi tersebut, tujuan penelitian dirumuskan yaitu menghasilkan data siap dianalisis lalu membandingkan kinerja algoritma Apriori dan FP-Growth dalam menemukan pola pembelian obat pada data transaksi Apotek Gretta Solok. Hasil analisis diharapkan dapat memberikan informasi yang berguna bagi pihak apotek dalam pengelolaan stok obat, strategi penataan produk, serta perencanaan yang lebih efektif dan berbasis data.

b. *Data Understanding* (Pemahaman Data)

Tahap kedua adalah mengumpulkan data awal dan melakukan eksplorasi untuk memahami karakteristik data. Data yang akan dianalisis merupakan data transaksi penjualan Apotek Gretta Solok dari bulan Oktober 2025 hingga Februari 2026 yang tersedia dalam bentuk *hardcopy* atau catatan manual. Keterbatasan ini terjadi karena data transaksi tidak terdokumentasi dengan baik akibat sistem pencatatan yang belum terkomputerisasi.

Pada tahap ini peneliti akan melakukan identifikasi awal terhadap beberapa aspek penting, yaitu:

1. Jenis informasi yang tercatat pada dokumen transaksi (tanggal, nama obat, jumlah)

2. Kelengkapan data transaksi yang tersedia,
3. Format pencatatan nama obat dan konsistensinya,
4. Potensi ketidak konsistenan penulisan akibat pencatatan manual, dan
5. Jumlah total transaksi yang tersedia untuk dianalisis.

Berdasarkan cara perolehan data tersebut, data yang digunakan dalam penelitian ini termasuk data primer. Hasil dari tahap ini akan menjadi dasar untuk menentukan langkah-langkah yang diperlukan pada proses persiapan data, termasuk mengidentifikasi tantangan yang mungkin muncul akibat karakteristik data manual dan menentukan strategi pembersihan data yang tepat.

c. *Data Preparation (Persiapan Data)*

Fase *data preparation* mencakup seluruh aktivitas pengolahan dan transformasi data mentah menjadi dataset yang siap digunakan dalam pemodelan. Karena data transaksi masih dalam bentuk pencatatan manual, maka pada tahap ini akan dilakukan beberapa langkah berikut:

1. Digitalisasi Data

Data transaksi yang masih tersedia dalam bentuk *hardcopy* dikonversi menjadi data digital dengan memanfaatkan kemampuan ekstraksi data berbasis *Large Language Model* (LLM), khususnya model multimodal *Vision-Language Model* (VLM) yang mampu membaca sekaligus memahami konteks dan struktur data pada gambar dokumen transaksi, termasuk tulisan tangan, sehingga hasil ekstraksi dapat langsung disusun ke dalam format Microsoft Excel untuk keperluan analisis. Meskipun ekstraksi data pada tahap ini memanfaatkan kemampuan LLM berbasis VLM tersebut, hasil ekstraksi tetap memerlukan verifikasi manual untuk memastikan kesesuaian data dengan dokumen asli. Tahap ini bertujuan untuk mempercepat proses digitalisasi data transaksi manual sekaligus menjaga akurasi data yang digunakan dalam penelitian.

2. Data Selection

Data selection merupakan tahap pemilihan atribut-atribut yang relevan dari data hasil digitalisasi untuk disesuaikan dengan kebutuhan analisis *association rule mining*. Pada tahap ini, atribut yang dipilih dan dipertahankan adalah tanggal transaksi sebagai dasar pembentukan identitas transaksi (*Transaction ID*) serta nama obat sebagai item utama dalam penemuan pola asosiasi, sedangkan atribut seperti jumlah barang (*quantity*) dan harga dieliminasi karena algoritma Apriori dan FP-Growth hanya memproses data representasi biner (keberadaan item dalam transaksi) dan bukan nilai numerik transaksi.

3. Pembersihan Data (*Data Cleaning*)

Data hasil digitalisasi kemudian melalui proses pembersihan untuk meningkatkan kualitas data. Proses ini meliputi:

- a. Mengkoreksi hasil ekstraksi untuk memastikan kesesuaian data dengan dokumen sumber.
- b. Menstandarisasi nama obat dengan memperbaiki kesalahan penulisan, menyeragamkan variasi penamaan, dan mengonversi singkatan ke bentuk baku (contoh: PCT dan Parasetamol menjadi Paracetamol).
- c. Menyaring *item* yang tidak termasuk kategori obat, seperti alat cek gula darah, kolesterol, dan asam urat.
- d. Menggabungkan dataset (menggabungkan seluruh *sheet* yang terpisah jadi satu halaman).
- e. Membuat identitas transaksi (*Transaction ID*) berdasarkan tanggal transaksi, dengan ketentuan satu hari dianggap sebagai satu transaksi (contoh: TRX-20250101).

4. Pembentukan Dataset ke Format *Basket*

Data yang sudah dibersihkan kemudian ditransformasikan ke dalam format transaksi (*transactional* atau *basket format*) sesuai format *association rule mining*. Setiap tanggal dianggap sebagai satu

transaksi dan seluruh obat yang tercatat pada tanggal tersebut dikelompokkan ke dalam transaksi yang sama. Hasil transformasi ini menghasilkan dataset yang siap digunakan untuk proses pencarian *frequent itemset* dan pembentukan aturan asosiasi menggunakan algoritma Apriori dan FP-Growth.

Dengan demikian, dataset yang dihasilkan telah valid dan siap digunakan dalam analisis menggunakan algoritma Apriori dan FP-Growth.

d. Modeling (Pemodelan)

Fase *modeling* merupakan tahap penerapan teknik analitik untuk mengekstraksi pola atau pengetahuan dari data yang telah dipersiapkan. Pada penelitian ini, proses pemodelan akan dilakukan dengan menerapkan dua algoritma *association rule mining*, yaitu Apriori dan FP-Growth, pada dataset yang sama sehingga hasilnya dapat dibandingkan secara objektif.

Langkah-langkah yang akan dilakukan pada tahap modeling adalah sebagai berikut:

1. Menentukan beberapa skenario nilai parameter yang akan diuji, meliputi variasi nilai minimum *support* dan minimum *confidence* untuk menemukan kombinasi parameter yang optimal.
2. Menerapkan algoritma Apriori pada dataset transaksi menggunakan *library mlxtend* pada bahasa pemrograman Python melalui Jupyter Notebook untuk setiap kombinasi parameter yang telah ditentukan.
3. Menerapkan algoritma FP-Growth pada dataset yang sama.
4. Menghasilkan aturan asosiasi dari masing-masing algoritma berdasarkan nilai *support* dan *confidence* yang ditetapkan.

Tahap ini bertujuan untuk memperoleh hasil pemodelan dari kedua algoritma yang nantinya akan dibandingkan pada tahap evaluasi berdasarkan kualitas aturan asosiasi yang dihasilkan dan efisiensi komputasinya.

e. Evaluation (Evaluasi)

Fase *evaluation* bertujuan menilai sejauh mana hasil pemodelan telah memenuhi tujuan *data mining* dan tujuan bisnis yang ditetapkan pada

tahap awal. Juga untuk menilai dan membandingkan kinerja algoritma Apriori dan FP-Growth berdasarkan beberapa kriteria yang telah ditetapkan. Evaluasi akan dilakukan melalui tiga aktivitas utama berikut:

1. Evaluasi Hasil Analisis

Membandingkan kualitas output dari kedua algoritma berdasarkan beberapa aspek, yaitu jumlah aturan asosiasi yang dihasilkan, rentang nilai *support* dan *confidence* dari aturan-aturan tersebut, nilai *lift ratio* yang menunjukkan kekuatan hubungan antara item, serta tingkat relevansi dan interpretabilitas aturan terhadap konteks bisnis Apotek Gretta Solok. Aturan yang dihasilkan akan divalidasi dengan pihak apotek untuk memastikan kesesuaiannya dengan kondisi di lapangan.

2. Evaluasi Efisiensi Komputasi

Mengukur performa teknis masing-masing algoritma melalui dua metrik utama, yaitu waktu eksekusi yang diukur menggunakan *library time* untuk mencatat durasi proses dari awal hingga menghasilkan aturan asosiasi final, serta pemantauan penggunaan memori menggunakan *memory_profiler* untuk mengetahui efisiensi penggunaan sumber daya komputasi selama proses berlangsung.

3. Analisis Perbandingan

Menyusun tabel komparasi yang merangkum seluruh proses untuk evaluasi dari kedua algoritma dan menganalisis perbedaan antara kualitas hasil analisis dan efisiensi komputasi. Analisis ini akan mempertimbangkan kelebihan dan kekurangan masing-masing algoritma untuk menentukan algoritma yang paling optimal diterapkan pada karakteristik data transaksi Apotek Gretta Solok.

Hasil dari tahap ini akan menjadi dasar dalam menentukan algoritma yang lebih efektif dan efisien untuk karakteristik data transaksi Apotek Gretta Solok, serta memberikan rekomendasi untuk penelitian atau implementasi selanjutnya.

f. *Deployment* (Penerapan)

Fase *deployment* merupakan tahap penerapan dan penyajian hasil *data mining* kepada pengguna akhir. Tahap ini akan dilakukan melalui tiga aktivitas utama berikut:

1. Dokumentasi Hasil

Mendokumentasikan seluruh proses analisis mulai dari persiapan data, implementasi algoritma, hingga evaluasi dalam bentuk *Jupyter Notebook* yang interaktif, lengkap dengan kode program, output hasil analisis, visualisasi data, serta penjelasan naratif untuk setiap tahapan agar dapat ditelusuri dan direproduksi kembali jika diperlukan di masa mendatang.

2. Penyajian Hasil

Menyusun laporan yang mudah dipahami oleh pihak non-teknis, meliputi ringkasan eksekutif temuan utama dengan bahasa sederhana, visualisasi pola pembelian dalam bentuk tabel atau diagram yang intuitif, daftar aturan asosiasi terkuat beserta interpretasi bisnisnya.

3. Rekomendasi Implementasi

Memberikan saran praktis berdasarkan hasil analisis, yaitu strategi penataan produk di rak dengan menempatkan obat-obat yang cenderung dibeli secara bersamaan di area yang berdekatan dalam kategorinya masing-masing guna meningkatkan kemudahan dan kenyamanan konsumen, serta optimalisasi pengelolaan stok dengan memprioritaskan ketersediaan produk-produk yang memiliki frekuensi pembelian tinggi berdasarkan pola transaksi sehingga risiko kekosongan stok dapat diminimalkan. Terakhir, usulan penerapan sistem pencatatan terkomputerisasi di masa mendatang untuk memudahkan analisis berkelanjutan dan meningkatkan kualitas data transaksi.

B. Setting Penelitian

1. Lokasi Penelitian

Penelitian ini dilaksanakan di Apotek Gretta Solok, yang berlokasi di Kampung Jawa, Kecamatan Tanjung Harapan, Kota Solok, Sumatera Barat. Pemilihan lokasi ini didasarkan pada pertimbangan bahwa Apotek Gretta Solok merupakan apotek yang memiliki 2 cabang dengan volume transaksi penjualan yang cukup tinggi, namun masih beroperasi secara konvensional. Kondisi ini menjadikan Apotek Gretta Solok sebagai objek penelitian yang tepat untuk penerapan algoritma *data mining*, khususnya dalam mengidentifikasi pola kombinasi pembelian obat yang dapat membantu meningkatkan efisiensi pengelolaan stok dan strategi penjualan. Selain itu, pihak Apotek Gretta Solok juga memberikan akses terhadap data transaksi penjualan yang diperlukan dalam penelitian ini serta bersedia untuk menjadi objek penelitian.

2. Jadwal Penelitian

Rangkaian perencanaan waktu penelitian dapat dilihat pada Tabel 3.1 penelitian selesai dalam waktu 8 bulan.

Tabel 3. 1 Jadwal Penelitian

Kegiatan	Bulan							
	1	2	3	4	5	6	7	8
Mengidentifikasi Masalah								
Studi Literatur								
<i>Business Understanding</i>								
<i>Data Understanding</i>								
<i>Data Preparation</i>								
<i>Modelling</i>								
<i>Evaluation</i>								
<i>Deployment</i>								

Dokumentasi								
-------------	--	--	--	--	--	--	--	--

C. Alat/Instrumen Penelitian

Adapun alat dan instrumen yang digunakan untuk menyelesaikan penelitian ini terdiri dari:

1. Perangkat Keras

Perangkat keras yang digunakan dalam penelitian ini adalah sebagai berikut:

- a. Processor AMD Ryzen 7 7435HS (16 CPUs), ~3.1GHz
- b. Memory (RAM) 8 GB
- c. GPU NVIDIA GeForce RTX 2050
- d. SSD 477 GB
- e. Mouse

2. Perangkat Lunak

Perangkat lunak yang digunakan dalam penelitian ini adalah sebagai berikut:

- a. Sistem Operasi Windows 11 Home Single Language 64-bit
- b. Jupyter Notebook
- c. Microsoft Office Word
- d. Microsoft Office Excel
- e. Browser (Microsoft Edge & Chrome)
- f. Bahasa pemrograman Python, dsb untuk penunjang penelitian.

D. Teknik Pengumpulan Data

Pada penelitian ini terdapat beberapa metode dalam pengumpulan data yang digunakan sebagai berikut :

1. Metode Observasi

Observasi merupakan metode pengumpulan data yang dilakukan melalui pengamatan langsung terhadap objek penelitian untuk memperoleh gambaran nyata mengenai kondisi di lapangan. Dalam penelitian ini, observasi difokuskan pada aktivitas operasional yang berkaitan dengan proses pencatatan transaksi penjualan dan pengelolaan

data di Apotek Gretta. Kegiatan observasi meliputi pengamatan terhadap mekanisme pencatatan penjualan, format data transaksi yang tersedia, serta cara penyimpanan data yang digunakan oleh pihak apotek. Observasi dilaksanakan secara langsung di Apotek Gretta yang berlokasi di Kampung Jawa, Kecamatan Tanjung Harapan, Kota Solok, Sumatera Barat.

2. Wawancara

Proses wawancara dilakukan dengan penanggung jawab Apotek Gretta Solok serta beberapa pegawai yang terlibat langsung dalam aktivitas penjualan untuk memperoleh informasi terkait mekanisme pencatatan transaksi dan pengelolaan data penjualan. Wawancara difokuskan pada beberapa aspek utama, yaitu cara pencatatan transaksi yang sedang berjalan, ketersediaan dan kelengkapan data penjualan, serta kendala yang dihadapi dalam pemanfaatan data transaksi sebagai dasar pengambilan keputusan. Adapun beberapa pertanyaan yang diajukan meliputi:

1. Bagaimana sistem penjualan yang digunakan di Apotek Gretta Solok, apakah sudah terkomputerisasi atau masih konvensional?
2. Bagaimana pengelolaan data penjualan yang dilakukan saat ini?
3. Apakah produk-produk dengan tingkat permintaan tinggi sering mengalami kehabisan stok?
4. Bagaimana cara penyusunan dan penataan barang di apotek saat ini?

Berdasarkan hasil wawancara tersebut, maka diperoleh informasi sebagai berikut:

1. Apotek Gretta Solok belum menggunakan sistem penjualan yang terkomputerisasi, semuanya masih beroperasi secara konvensional.
2. Data penjualan hanya digunakan sebagai arsip dan belum dimanfaatkan untuk menganalisis pola pembelian guna mendukung strategi bisnis.
3. Beberapa produk dengan permintaan tinggi memang sering habis, hal ini disebabkan karena kurangnya evaluasi secara berkala terkait penjualan obat.

4. Barang-barang di Apotek juga belum dikelompokkan berdasarkan kategori.

Informasi yang diperoleh dari wawancara ini digunakan untuk memahami kondisi nyata di lapangan dan sebagai dasar dalam merancang proses analisis data menggunakan teknik *data mining*.

3. Metode Studi Pustaka

Studi pustaka merupakan metode pengumpulan data yang dilakukan melalui penelusuran dan kajian terhadap berbagai sumber literatur yang relevan dengan topik penelitian, seperti buku, jurnal ilmiah dan skripsi terdahulu. Metode ini digunakan untuk memperoleh landasan teori, konsep, serta hasil penelitian sebelumnya yang mendukung analisis dan pemecahan masalah dalam penelitian.

E. Teknik Analisis Data

Penelitian ini menggunakan dua algoritma *association rule mining*, yaitu algoritma Apriori dan algoritma FP-Growth untuk menganalisis pola pembelian obat pada data transaksi penjualan Apotek Greta Solok. Berikut adalah penjelasan teknik analisis data yang digunakan:

1. Langkah Analisis Menggunakan Algoritma Apriori

Algoritma Apriori merupakan metode untuk menemukan pola pembelian barang yang sering dibeli bersamaan dengan menggunakan prinsip "jika suatu kelompok barang sering muncul bersama, maka bagian-bagian kecil dari kelompok tersebut juga pasti sering muncul dan sebaliknya, jika suatu kelompok barang jarang muncul, maka kelompok yang lebih besar yang mengandungnya juga pasti jarang muncul" (Han dkk., 2012). Langkah-langkah penerapan algoritma Apriori dalam penelitian ini adalah (Han dkk., 2012):

- a. Menentukan nilai batas *minimum support* dan *confidence*.
- b. Menghitung frekuensi kemunculan setiap item.
- c. Mengeliminasi item yang tidak memenuhi batas minimal kemunculan.
- d. Membentuk kombinasi dua item, lalu menghitung frekuensinya.

- e. Mengeliminasi kombinasi yang tidak memenuhi batas.
- f. Mengulangi pembentukan kombinasi hingga tidak ada kombinasi baru yang memenuhi syarat.
- g. Membentuk aturan hubungan antara *item* dan menghitung *confidence*.
- h. Menyaring aturan yang memenuhi batas *minimum confidence*.

2. Langkah Analisis Menggunakan Algoritma Fp-Growth

Algoritma FP-Growth merupakan metode untuk mining *frequent itemset* tanpa melakukan *candidate generation* (membentuk banyak kandidat kombinasi produk) dengan menggunakan struktur data FP-Tree (*Frequent Pattern Tree*) yang lebih efisien (Rustam dkk., 2024). Langkah-langkah penerapan algoritma FP-Growth dalam penelitian ini adalah sebagai berikut (Han dkk., 2012):

- a. Menentukan nilai batas minimal *support* dan *confidence*.
- b. Menghitung frekuensi setiap item.
- c. Mengurutkan item berdasarkan frekuensi tertinggi.
- d. Membentuk struktur pohon frekuensi (FP-Tree) berdasarkan transaksi.
- e. Membentuk pola kombinasi item dari FP-Tree tersebut.
- f. Menghasilkan kombinasi yang memenuhi batas *minimum support*.
- g. Membentuk aturan hubungan dan menghitung *confidence*.
- h. Menyaring aturan sesuai batas yang ditentukan.