

BAB 3

METODE PENELITIAN

Penelitian ini menggunakan pendekatan studi kasus dengan mengikuti tahapan *Knowledge Discovery in Databases* (KDD) untuk menggali pola-pola tersembunyi dalam data. Proses dimulai dari tahap pemilihan data (*data selection*), pembersihan data (*data cleaning*), hingga transformasi data ke dalam format yang sesuai untuk dianalisis.

Selanjutnya, algoritma *Frequent Pattern Growth* (FP-Growth) digunakan untuk menemukan pola asosiasi antar item yang sering muncul bersama. Pola-pola yang dihasilkan kemudian dievaluasi untuk memastikan validitas dan relevansinya. Hasil evaluasi tersebut diimplementasikan sebagai rekomendasi fitur tambahan dalam sistem perpustakaan yang telah ada di Perpustakaan Daerah Soekarno-Hatta Kabupaten Tegal

3.1 BAHAN DAN ALAT PENELITIAN

Bahan penelitian ini terdiri dari data primer dan data sekunder. Data primer berupa salinan data peminjaman buku dari Perpustakaan Daerah Soekarno-Hatta Kabupaten Tegal yang mencakup periode tahun 2012 hingga 2025. Data ini berisi informasi seperti ID anggota, judul buku, nama pengarang, nomor klasifikasi buku, dan tanggal peminjaman. Jumlah data peminjaman buku yang digunakan dalam penelitian ini lebih dari seribu baris.

Untuk memperkaya informasi, data klasifikasi buku berdasarkan sistem *Dewey Decimal Classification* (DDC) digunakan. Data klasifikasi buku mencakup kode klasifikasi dari 000 hingga 999 beserta deskripsi masing-masing kategori.

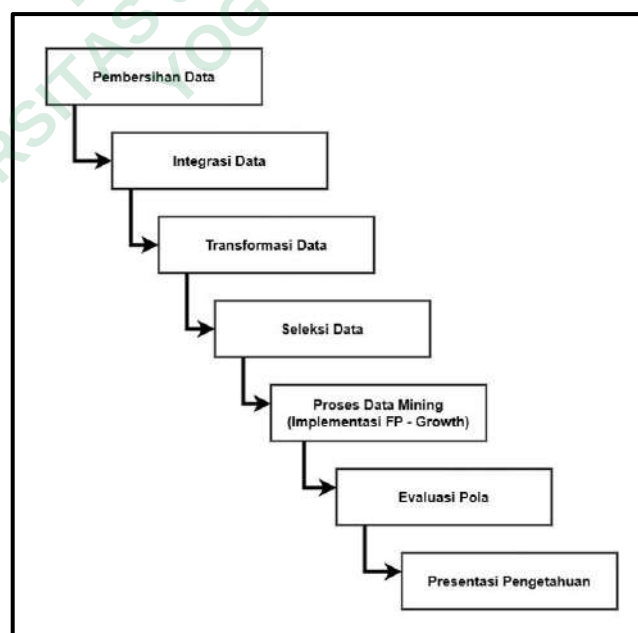
Sementara itu, data sekunder diperoleh dari literatur berupa jurnal dan buku yang membahas algoritma FP-Growth, teknik asosiasi dalam *data mining*, serta evaluasi pola asosiasi menggunakan metrik *lift*, *leverage*, *conviction* dan *fisher's exact test*. Data sekunder ini digunakan sebagai landasan teoritis dan acuan dalam pengujian pola yang dihasilkan.

Perangkat keras yang digunakan dalam penelitian ini adalah laptop dengan spesifikasi yang mendukung pengoperasian sistem serta perangkat lunak yang diperlukan dengan konektivitas internet yang stabil. Sistem operasi serta perangkat lunak yang digunakan dalam penelitian ini mencakup:

1. Sistem Operasi: Windows 11 23H2
2. Text Editor: Visual Studio Code 1.198
3. Bahasa Pemrograman: Python 3.13.2
4. Basis Data: MariaDB 10.4.32
5. Pustaka: Flask v3.1.1, Jupyter Notebook v6.5.4, Matplotlib v3.7.1, Mlxtend 0.23.4, Numpy v2.3.0, Openpyxl v3.1.5, Pandas v1.5.3, Scikit-learn v1.7.0, Scipy v1.15.3, SQLAlchemy, 2.0.41, Statsmodels v0.14.0.

3.2 JALAN PENELITIAN

Penelitian ini menggunakan pendekatan *Knowledge Discovery in Databases* (KDD) sebagai langkah utama untuk mengolah dan menganalisis data peminjaman buku di perpustakaan. Metode KDD dipilih karena mampu membantu menemukan informasi penting dari data yang banyak dan beragam secara terstruktur dan sistematis.



Gambar 3.1 Metode KDD

Data yang awalnya belum teratur dan mungkin tidak lengkap akan dibersihkan, digabungkan, dan diubah menjadi bentuk yang siap untuk dianalisis. Setelah itu, algoritma *Frequent Pattern Growth* (FP-Growth) digunakan untuk mencari pola-pola buku yang sering dipinjam bersama. Hasil analisis tersebut kemudian dievaluasi untuk memastikan kualitasnya sebelum akhirnya disajikan dalam bentuk sistem sederhana yang mudah digunakan. Adapun tahapan dalam proses KDD yang dilakukan dalam penelitian ini sebagai berikut:

1. Pembersihan Data (*Data Cleaning*)

Penelitian ini dimulai dengan tahap pembersihan data pada *dataset* peminjaman buku yang diperoleh dari staf perpustakaan. Tujuan dari tahap ini adalah untuk menghapus data yang tidak konsisten, mengandung *noise*, tidak relevan, maupun tidak digunakan dalam analisis.

2. Integrasi Data (*Data Integration*)

Setelah data dibersihkan, tahap integrasi dilakukan dengan menggabungkan data transaksi peminjaman dan data koleksi buku berdasarkan nomor klasifikasi. Proses ini bertujuan untuk menambahkan informasi pada data pada data transaksi peminjaman dengan data klasifikasi buku. Sebelum penggabungan, format data diseragamkan agar tidak terjadi kesalahan dalam proses integrasi.

3. Transformasi Data (*Data Transformation*)

Pada tahap ini, data yang telah terintegrasi diubah strukturnya agar sesuai dengan kebutuhan proses analisis pola asosiasi. Transformasi dilakukan dengan cara mengelompokkan data berdasarkan ID anggota dan tanggal peminjaman, sehingga setiap transaksi mewakili satu sesi peminjaman yang terdiri dari beberapa judul buku. Hasil dari tahap ini adalah *dataset* dalam bentuk transaksi per anggota per hari.

4. Seleksi Data (*Data Selection*)

Setelah data ditransformasikan, dilakukan proses seleksi untuk menyaring data yang relevan dan memenuhi kriteria analisis. Data yang tidak lengkap, duplikat, atau transaksi dengan hanya satu buku dihapus karena tidak dapat

membentuk asosiasi yang bermakna. Dataset hasil seleksi ini kemudian digunakan dalam proses *data mining*.

5. Proses *Data mining*

Setelah data diubah kedalam bentuk yang sesuai, proses *data mining* dilakukan dengan menggunakan algoritma *Frequent Pattern Growth (FP-Growth)* untuk menemukan pola asosiasi antar kategori buku yang sering dipinjam secara bersamaan. Proses ini terdiri dari beberapa langkah sebagai berikut:

a. Penetapan Nilai *Minimum Support* Berdasarkan Rata-rata

Sebelum menjalankan algoritma FP-Growth, dilakukan analisis frekuensi kemunculan kombinasi kategori buku dalam data transaksi. Nilai *minimum support* ditentukan dengan mengambil rata-rata geometrik dari frekuensi kemunculan item pada seluruh data transaksi, sehingga hanya item dengan kemunculan yang memiliki frekuensi diatas rata – rata yang akan diproses lebih lanjut.

b. Penerapan Algoritma FP - Growth

Algoritma FP-Growth kemudian digunakan untuk menghasilkan *frequent itemset*, yaitu kombinasi dua atau lebih kategori buku yang memenuhi ambang *support* tersebut.

c. Pembentukan Aturan Asosiasi (*Association Rules*)

Dari *frequent itemset* yang diperoleh, dibentuk aturan asosiasi untuk mengidentifikasi keterkaitan antar kategori buku. Misalnya, apabila kategori A sering dipinjam bersamaan dengan kategori B, maka akan terbentuk aturan $A \Rightarrow B$.

6. Evaluasi Pola (*Pattern Evaluation*)

Setelah pembentukan aturan asosiasi berhasil, langkah berikutnya adalah melakukan evaluasi terhadap pola-pola asosiasi yang terbentuk. Tujuan dari tahap ini adalah untuk menilai sejauh mana kekuatan dan signifikansi hubungan antar kategori buku yang sering dipinjam secara bersamaan, serta menyaring pola yang benar-benar bermakna. Evaluasi dilakukan dengan menggunakan sejumlah metrik, antara lain:

a. *Lift*

Lift mengukur kekuatan asosiasi antara dua item dengan membandingkan probabilitas kemunculan bersamaan dengan probabilitas kemunculan masing-masing item secara independen. Nilai *lift* > 1 menunjukkan bahwa dua item lebih mungkin muncul bersama daripada secara acak, yang mengindikasikan adanya asosiasi positif yang kuat.

b. *Leverage*

Leverage menghitung selisih antara probabilitas aktual dua item muncul bersama dengan probabilitas yang diharapkan jika keduanya independen. Semakin besar nilai *leverage*, semakin kuat asosiasi tersebut, karena menunjukkan bahwa kemunculan bersama tersebut jauh dari sekadar kebetulan.

c. *Conviction*

Conviction mengukur tingkat keandalan suatu aturan asosiasi dengan mempertimbangkan kemungkinan terjadinya *consequent* tanpa adanya *antecedent*. Nilai *conviction* yang lebih tinggi menunjukkan bahwa kehadiran *antecedent* benar-benar memperbesar peluang munculnya *consequent*.

d. *Fisher Exact Test*

Untuk memastikan pola yang terbentuk benar-benar signifikan secara statistik, dilakukan uji validasi menggunakan *Fisher Exact Test*. Uji ini digunakan untuk mengevaluasi kekuatan hubungan antara dua item dalam kontingensi 2×2 , terutama ketika jumlah data tidak terlalu besar. Nilai *p-value* yang lebih kecil dari ambang batas tertentu menunjukkan bahwa asosiasi yang ditemukan signifikan secara statistik dan tidak terjadi secara acak.

Seluruh metrik tersebut diterapkan secara berurutan untuk mengevaluasi setiap aturan asosiasi yang terbentuk. Pola-pola dengan nilai evaluasi yang rendah akan dieliminasi, sedangkan pola dengan nilai yang tinggi akan dipertahankan sebagai hasil akhir. Proses ini memastikan bahwa hanya pola

yang kuat, signifikan, dan relevan yang digunakan sebagai dasar untuk pengambilan keputusan.

7. Presentasi Pengetahuan (*Knowledge Presentation*)

Tahap akhir dari proses ini adalah menyajikan hasil analisis dalam bentuk sistem sederhana. Sistem ini berfungsi sebagai media untuk menyampaikan pengetahuan yang telah ditemukan kepada pengguna dalam bentuk yang mudah dipahami. Pendekatan ini memungkinkan hasil analisis tidak hanya tersimpan sebagai laporan, tetapi juga dapat dimanfaatkan langsung melalui antarmuka sistem yang interaktif.

PERPUSTAKAAN
UNIVERSITAS JENDERAL ACHMAD YANIL
YOGYAKARTA