

BAB 3

METODE PENELITIAN

Penelitian penulis merupakan kategori penelitian kuantitatif eksploratif dengan menggunakan Algoritma Apriori dan metode *Knowledge Discovery in Databases* untuk mengidentifikasi pola buku yang dipinjam secara bersamaan di Dinas Perpustakaan dan Kearsipan Kabupaten Sukoharjo, yang terdiri dari beberapa tahapan sebagai berikut:

1. *Selection* (Seleksi Data)

Memperoleh data dari sistem layanan sirkulasi di Dinas Perpustakaan dan Kearsipan Kabupaten Sukoharjo yaitu Inlislite. Kemudian data dalam beberapa tahun dijadikan ke dalam satu dataset yang memiliki atribut seperti nomor anggota, tanggal pinjam, tanggal kembali, data bibliografis, nama anggota, jenis kelamin, nama petugas, dan lain sebagainya. Melakukan pengecekan pada dataset yang sudah terkumpul dengan melakukan penyesuaian pada atributnya.

2. *Preprocessing*

Membersihkan dataset untuk mempermudah proses analisis dengan melakukan beberapa tahapan seperti menghapus dataset yang memiliki data kosong atau tidak lengkap, menghilangkan duplikasi pada data serta melakukan pengurangan data yang tidak sesuai atau berkaitan dengan tujuan analisis

3. *Transformation*

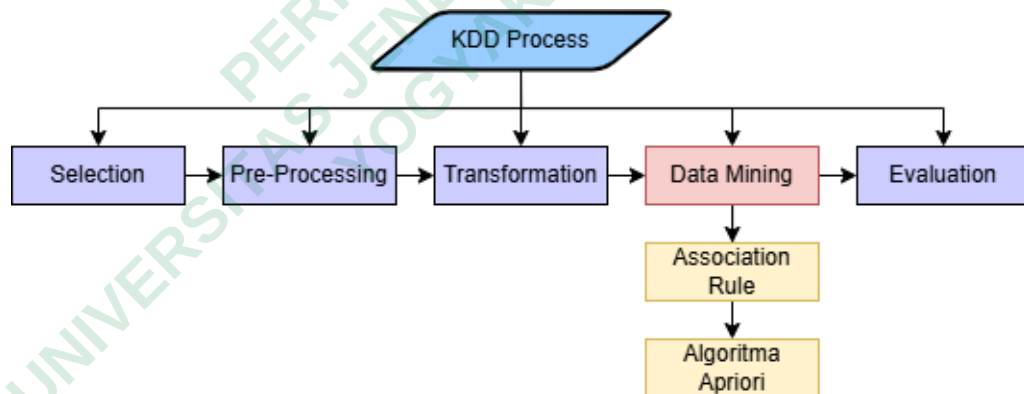
Melakukan konversi pada dataset menjadi bentuk transaksi itemset (biner 1 dan 0) yang menunjukkan kombinasi buku yang sering dipinjam bersamaan. Pada tahap ini pengelompokan dataset berdasarkan nomor anggota dan judul buku sehingga menghasilkan data peminjaman buku dalam bentuk transaksi yang siap dilakukan penerapan data mining.

4. Data Mining

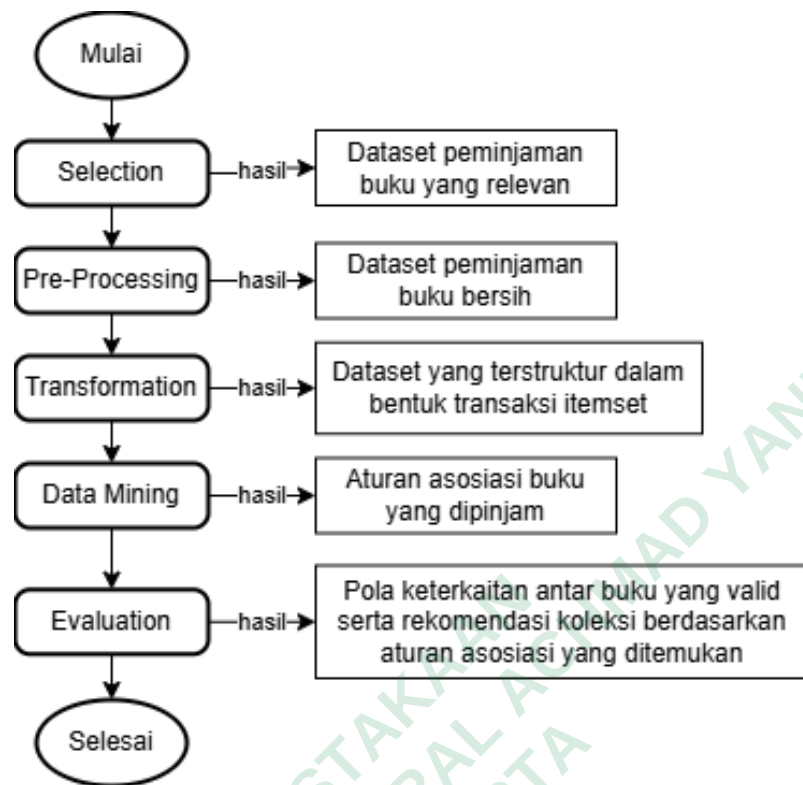
Menganalisis pola buku yang sering dipinjam bersamaan dengan menerapkan Algoritma Apriori untuk mengidentifikasi pola keterkaitan (*Association Rule*) pada data peminjaman buku, dengan melihat dua parameter utama yaitu *support* (kemunculan kombinasi buku dalam dataset transaksi peminjamna dan *confidence* (peluang anggota meminjam buku yang berkaitan).

5. Evaluation

Mengevaluasi hasil pengolahan dataset pada proses data mining untuk menentukan aturan dengan pola yang konsisten. Evaluasi dilakukan dengan melakukan perbandingan pada nilai *support* dan *confidence*, kemudian setelah menemukan nilai yang optimal, perlu dilakukan visualisasi untuk mempermudah penyampaian informasi hasil penelitian yang dimanfaatkan dalam membantu pengambilan keputusan di perpustakaan Kabupaten Sukoharjo.



Gambar 3.1 Metode Penelitian



Gambar 3.2 Tahapan KDD

3.1 BAHAN DAN ALAT PENELITIAN

Bahan penelitian berupa data peminjaman buku bulan Januari – Desember tahun 2021 – 2024 di Dinas Perpustakaan dan Kearsipan Kabupaten Sukoharjo

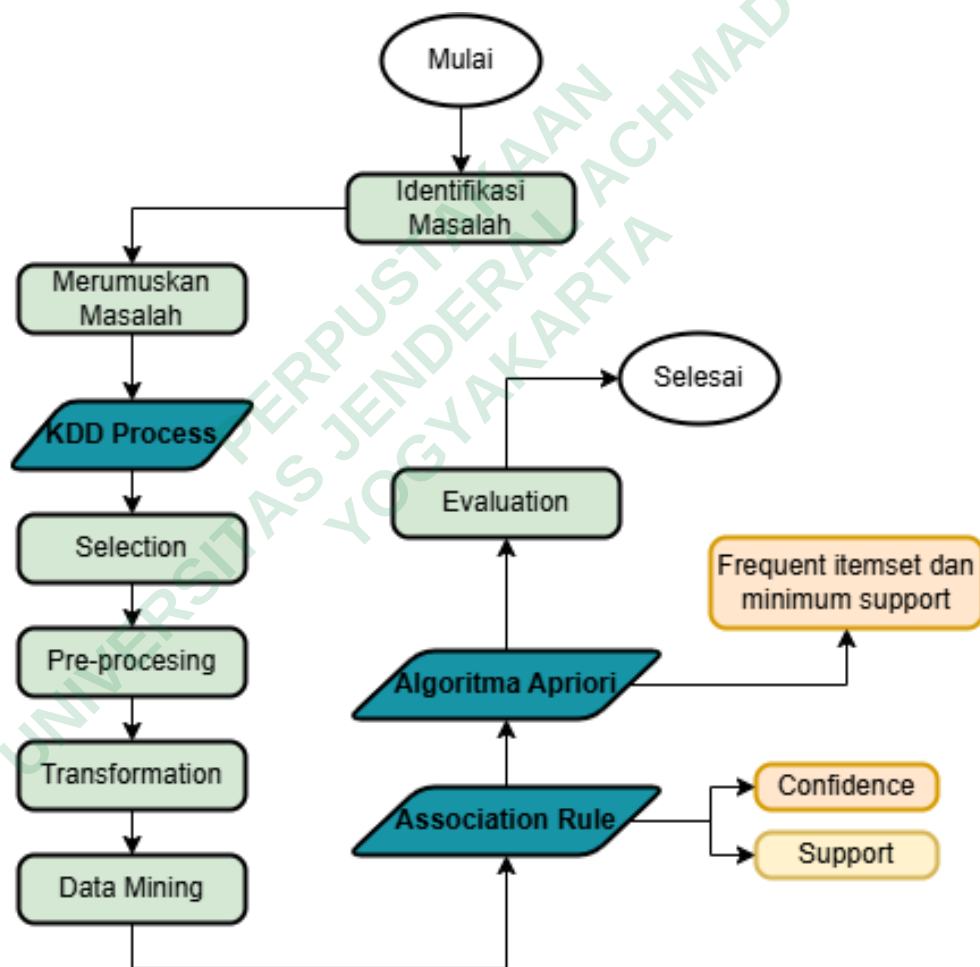
Alat yang digunakan dalam melakukan penelitian menggunakan laptop dengan spesifikasi yang memadai sehingga sistem operasi dan *software* analisis dapat digunakan dengan optimal, dan memiliki koneksi internet yang stabil.

Sistem operasi dan aplikasi yang digunakan dalam penelitian ini meliputi:

1. Sistem Operasi: Windows 10
2. Memory: 4 GB RAM
3. Software Analisis: *Google Colaboratory* (implementasi algoritma apriori), Python (termasuk pustaka seperti Pandas, Mlxtend, dan Matplotlib) dan *Microsoft Excel*.

3.2 JALAN PENELITIAN

Penelitian terkait analisis keterkaitan buku dari data peminjaman buku di Dinas Perpustakaan dan Kearsipan Kabupaten Sukoharjo menggunakan metode *Knowledge Discovery in Database (KDD)* pada pemodelan *association rule* untuk pengolahan data dan jalan penelitian ini akan dilakukan berdasarkan tahapan dari metode tersebut. Metode ini dipilih karena memberikan fokus pada kebutuhan pengguna dan permasalahan yang sedang dialami, selain itu metode KDD memiliki proses yang berulang untuk memahami secara mendalam kebutuhan dan preferensi data sehingga dapat memberikan solusi inovatif dan sesuai dengan kebutuhan.



Gambar 3.3 Jalan Penelitian

Berikut merupakan penjelasan terkait jalan penelitian pada Gambar 3.3:

1. Identifikasi dan Analisis Masalah

Melakukan analisis masalah dengan melihat kondisi di Dinas Perpustakaan dan Kearsipan Kabupaten Sukoharjo yang memiliki jumlah koleksi sebanyak 56.067 eksemplar yang mencakup 32.750 judul. Berdasarkan data statistik tahun 2024 jumlah pengunjung mencapai 10.191 dengan total peminjaman buku sebesar 3.445. Berdasarkan data tersebut, perpustakaan Kabupaten Sukoharjo memiliki jumlah pengunjung dan koleksi yang tinggi serta data peminjaman buku yang sudah tercatat secara digital, tetapi data peminjaman buku belum diolah menjadi informasi yang bermanfaat.

Pengunjung masih kesulitan dalam menemukan buku yang akan dipinjam dikarenakan kurangnya preferensi dari pustakawan. Hal itu dikarenakan belum ada cara terstruktur yang digunakan oleh perpustakaan Kabupaten Sukoharjo untuk mengetahui buku yang sering dipinjam dan belum terdapat strategi penyusunan rak berdasarkan hasil analisis pola data peminjaman buku.

2. Merumuskan Masalah

Berdasarkan identifikasi masalah, Perpustakaan Kabupaten Sukoharjo pwelu menerapkan data mining untuk mengelola data peminjaman buku. Salah satu data mining yang dapat diterapkan adalah *asosiation rule* dengan algoritma apriori dan metode *Knowledge Discovery in Databases*. Hal tersebut bertujuan untuk menemukan pola keterkaitan antar buku yang sering dipinjam berdasarkan pada data peminjaman buku di Dinas Perpustakaan dan Kearsipan Kabupaten Sukoharjo. Hasil dari penerapan data mining pada data peminjaman buku pada Dinas Perpustakaan dan Kearsipan Kabupaten Sukoharjo dapat meningkatkan kualitas pelayanan dan mendukung pengelolaan koleksi di Perpustakaan Kabupaten Sukoharjo.

3. Selection (Pengambilan Data)

Data yang digunakan dalam penelitian ini berasal dari sistem Aplikasi Inlislite yang diterapkan di Dinas Perpustakaan dan Kearsipan Kabupaten Sukoharjo untuk melakukan proses sirkulasi perpustakaan. Dataset memiliki beberapa atribut seperti tanggal pinjam, tanggal jatuh tempo, tanggal kembali,

jumlah hari telat, nomor induk, data bibliografis, nomor kelas, nomor anggota, nama anggota, nama anggota, jenis kelamin, kelompok umur, nama petugas dan nama petugas pengembalian. Berdasarkan penerapan algoritma apriori, tidak semua atribut digunakan untuk proses analisis data, cukup atribut nomor anggota dan data bibliografis.

Atribut nomor anggota menjadi identitas unik dari masing-masing anggota perpustakaan Kabupaten Sukoharjo. Selain itu, nomor anggota juga menunjukkan siapa yang melakukan peminjaman untuk menganalisis pola peminjaman. Atribut data bibliografis menunjukkan judul buku dan identitas buku seperti penerbit, edisi cetakan dan tempat publikasi. Proses ini memastikan bahwa dataset sudah memiliki atribut yang sesuai untuk proses analisis dengan menerapkan algoritma apriori dan *Knowledge Discovery in Databases*.

4. *Pre-Processing* (Pembersihan Data)

Proses pembersihan data dilakukan untuk memastikan bahwa dataset yang digunakan dalam proses analisis sudah bersih dan tidak memiliki kesalahan yang dapat mempengaruhi cara kerja algoritma dan hasil analisisnya. Beberapa proses yang dapat dilakukan pada tahap *pre-processing* seperti melakukan pengecekan data *null* (kosong), kemudian menghapus data duplikat, melakukan penanganan jika terdapat *missing value* dan jika diperlukan melakukan penyesuaian format pada atribut yang belum sesuai dengan proses analisis dan penerapan algoritma.

5. Transformasi Data

Proses transformasi data dilakukan setelah dataset sudah dipastikan bersih dan siap untuk dilakukan analisis. Pada tahap ini mengubah dataset menjadi format data transaksi yang disesuaikan dengan penerapan algoritma apriori. Tahapan transformasi mengubah dataset menjadi bilangan biner 1 dan 0, dimana setiap transaksi dikelompokkan berdasarkan nomor anggota dan judul buku yang dipinjam dalam satu baris. Transformasi perlu dilakukan untuk memastikan bahwa dataset bisa diterapkan teknik data mining *association rule*, sehingga mendapatkan hasil analisis yang optimal.

6. Data Mining (Penerapan *Association Rule* dan Algoritma Apriori)

Pada tahap data mining mulai menerapkan *association rule* dengan algoritma apriori untuk menemukan pola hubungan buku pada data peminjaman buku di Perpustakaan Kabupaten Sukoharjo dengan menggunakan *tools Google Collaboratory* untuk mengolah data secara efisien. Pentingnya pola buku dapat dilihat berdasarkan nilai *support* (nilai penunjang) merupakan presentase dari pola yang terdapat dalam dataset dan *confidence* (nilai kepastian) menunjukkan kekuatan hubungan antar item pada pola asosiasi (Fitri Polorida Ginting & Saripurna, 2021).

Berikut merupakan tahapan untuk membentuk hubungan judul buku yaitu:

a. Menganalisis pola yang memiliki frekuensi tinggi

Pada tahap ini menemukan pola item yang memenuhi syarat minimum dari nilai *support* pada dataset transaksi. Perhitungan nilai *support* dari suatu item dapat dirujuk pada rumus 1 berikut:

$$Support(A) = \frac{Jumlah\ Transaksi\ Mengandung\ A}{Total\ Transaksi} \quad (1)$$

sedangkan perhitungan untuk nilai *support* dari dua item dapat ditemukan pada rumus 2 berikut: (Fitri Polorida Ginting & Saripurna, 2021).

$$Support(A \cap B) = \frac{Jumlah\ Transaksi\ Mengandung\ A\ dan\ B}{Total\ Transaksi} \quad (2)$$

b. Membentuk aturan asosiasi

Pada tahap ini dilakukan jika sudah mendapatkan pola yang memiliki frekuensi tinggi, kemudian dilakukan pencarian aturan asosiasi yang memenuhi syarat minimum untuk *confidence* aturan asosiasi. Aturan *confidence* dinyatakan dalam bentuk "Jika A, maka B" dan dapat dilihat pada rumus 3 berikut: (Ibezato Zalukhu et al., 2024).

$$Confidence = P(B|A) = \frac{Jumlah\ Transaksi\ Mengandung\ A\ dan\ B}{Jumlah\ Transaksi\ Mengandung\ A} \quad (3)$$

7. Evaluasi

Pada tahap ini hasil dari proses data mining dibentuk kesimpulan dengan pola mana yang memiliki frekuensi tinggi untuk menjadi dasar pengambilan keputusan di Perpustakaan Kabupaten Sukoharjo. Analisis data berbentuk rekomendasi yang berisi judul buku dengan mencantumkan nilai *support* dan *confidence* dari setiap pola transaksi. Dalam tahap evaluasi juga diperlukan visualisasi untuk menyajikan hasil dalam bentuk visual untuk mempermudah pemahaman. Berikut adalah beberapa metode yang digunakan dalam visualisasi data:

a. Visualisasi tabel

Hasil dari pola peminjaman buku dimasukkan dalam tabel yang menampilkan aturan asosiasi antar buku berdasarkan nilai *support* dan *confidence*.

b. Diagram jaringan (*Network Graph*)

Visualisasi ini berbentuk grafik jaringan yang menunjukkan pola pada buku yang sering dipinjam. Diagram ini menggambarkan hubungan antara buku yang sering dipinjam. *Node* merepresentasikan judul buku dan *edge* (garis penghubung) menunjukkan hubungan peminjaman antara dua buku, dimana jika garis penghubungnya (*edge*) semakin tebal, maka hubungan pola antar buku semakin kuat.

c. *Heatmap*

Heatmap digunakan untuk menggambarkan seberapa sering dua buku dipinjam dalam suatu transaksi. Warna dalam *heatmap* menunjukkan tingkat kekuatan (*confidence*) dari hasil analisis pola antar buku, yang mana semakin kuat hubungan buku, warnanya juga lebih gelap.