

BAB III

METODE PENELITIAN

Metode penelitian dalam data mining merujuk pada serangkaian langkah atau prosedur yang digunakan untuk mengekstraksi informasi yang berharga dari sekumpulan data yang besar. Data mining adalah proses penggalian pola, penemuan informasi tersembunyi, dan analisis yang dilakukan pada data yang besar dan kompleks untuk mengidentifikasi hubungan, tren, dan pola yang dapat digunakan untuk pengambilan keputusan.

Beberapa metode penelitian yang digunakan dalam data mining meliputi:

- Klasifikasi: Mengelompokkan data ke dalam kategori atau kelas berdasarkan fitur atau karakteristik tertentu.
- Clustering (Pengelompokan): Memisahkan data menjadi kelompok atau cluster berdasarkan kesamaan karakteristik.
- Regresi: Menganalisis hubungan antara variabel untuk memahami bagaimana satu variabel memengaruhi yang lain.
- Asosiasi: Menemukan hubungan dan ketergantungan antara variabel dalam data.

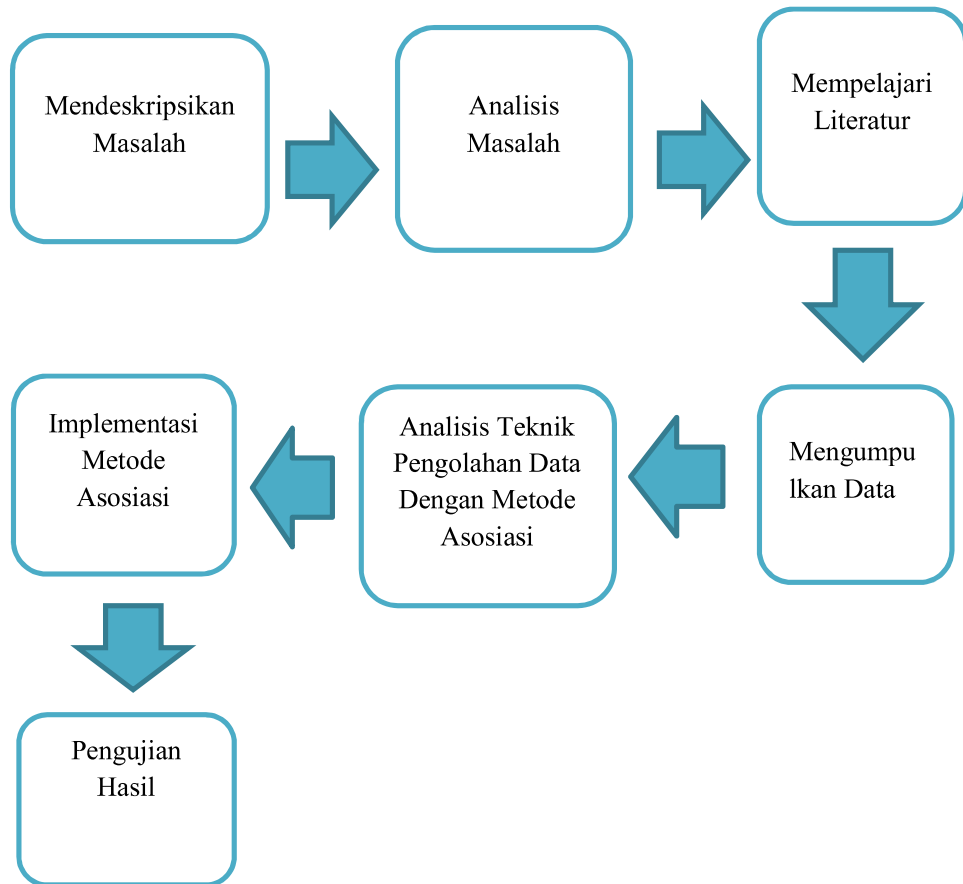
3.1. Desain Penelitian

Penelitian ini memanfaatkan metode asosiasi, suatu metode data mining yang digunakan untuk menemukan pola frekuensi tinggi dalam basis data. Pola frekuensi tinggi merujuk pada pola-pola item yang muncul secara sering dalam basis data, dengan frekuensi atau dukungan melebihi batasan tertentu yang disebut minimum support. Algoritma apriori, sebagai metode yang digunakan, hanya memerlukan

frequent itemset yang sudah diketahui sebelumnya untuk melanjutkan pemrosesan informasi. Proses pembentukan kombinasi itemsets dan pembuatan aturan dapat dijelaskan sebagai berikut(D. M. Sinaga et al., 2022) :

- a) Tahap awal melibatkan pembentukan kandidat itemset, yang dihasilkan dari kombinasi (k-1)-itemset yang telah diperoleh pada iterasi sebelumnya. Dalam algoritma apriori, satu pendekatan adalah melakukan pemangkasan pada kandidat k-itemset dengan menghilangkan subset yang terdiri dari k-1 item yang tidak terdapat dalam pola frekuensi tinggi dengan panjang k-1.
- b) Kemudian, langkah selanjutnya adalah menghitung support untuk setiap kandidat k-itemset. Support dari setiap kandidat ini dihitung dengan melakukan pemindaian pada basis data untuk menentukan jumlah transaksi yang mengandung seluruh item dalam kandidat k-itemset tersebut. Pendekatan ini merupakan bagian dari algoritma apriori yang memerlukan perhitungan dengan melakukan pemindaian pada seluruh basis data untuk k-itemset terpanjang.
- c) Pola frekuensi tinggi ditentukan dengan menetapkan kandidat k-itemset yang memiliki support lebih besar dari minimum support.
- d) Jika tidak ada pola frekuensi tinggi yang ditemukan, maka seluruh proses dihentikan. Namun, jika ada, nilai k akan ditambah satu dan algoritma kembali ke langkah pertama.

Desain penelitian dapat digambarkan pada gambar berikut :



(Sumber : (D. F. S. Sinaga, 2021))

Adapun desain dalam pengkajian ini sebagai berikut, yaitu:

1) Mendeskripsikan Masalah

Penjelasan pada pendekatan pertama ini bermula dari analisis. Memahami dengan mendalam masalah yang akan diselidiki, menetapkan cakupan dan batasan dari dilema tersebut sehingga membantu dalam pemahaman terhadap solusi potensial terkait dengan masalah yang tengah diteliti. Dalam fase analisis

ini, deskripsi masalah dijalankan melalui survei dan eksplorasi terhadap kendala yang terdapat pada objek yang sedang dianalisis.

2) Analisis Masalah

Langkah selanjutnya dalam proses penelitian melibatkan analisis permasalahan yang ditemukan dari hasil survei. Selanjutnya, permasalahan tersebut akan dikelompokkan berdasarkan situasi kritis, memberikan peluang bagi peneliti untuk mengidentifikasi solusi dan potensi perbaikan terhadap permasalahan yang ada.

3) Studi Literatur

Dalam upaya menyelesaikan penulisan ini, pemeriksa telah menggali literatur terkait dengan konsep penerapan aturan asosiasi dan algoritma pemecah apriori. Sumber referensi literatur dapat ditemukan di buku, jurnal, dan situs web, dan dapat berperan sebagai dukungan tambahan saat menentukan pendekatan yang cocok untuk permasalahan yang sedang diteliti.

4) Pengumpulan Data

Tahapan selanjutnya dalam pengumpulan data melibatkan observasi eksklusif dan wawancara dengan PT Excellindo Perkasa, sehingga permasalahan yang dihadapi dapat teridentifikasi dengan jelas. Setelah itu, peneliti melakukan penelitian kepustakaan dengan cara membaca buku dan jurnal yang relevan dengan penelitian saat ini. Data yang digunakan dalam penelitian ini merupakan data transaksi sisi pembelian.

5) Seleksi Data

Data yang dihasilkan dari proses pengumpulan data kemudian mengalami tahap seleksi, di mana data yang telah dipilih disaring untuk menghindari duplikasi, memperbaiki kesalahan tipografi, dan menghapus data yang kosong. Pada tahap ini, data yang telah dipilih akan digunakan dalam proses penambahan data.

6) Pengolahan Data

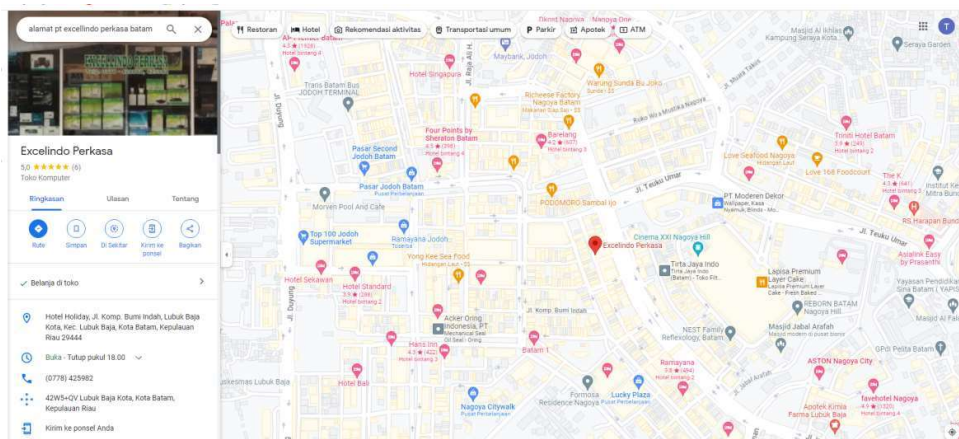
Pada fase ini, data yang telah dipilih dalam tahap seleksi diolah menggunakan metode aturan asosiasi-prioritas untuk mendapatkan nilai *confidence* dan *support* yang kemudian diterapkan untuk meningkatkan penjualan.

7) Pengujian Hasil

Menjalankan serangkaian prosedur dalam penelitian tentu akan membawa kita ke arah hasil yang diinginkan oleh peneliti. Pada tahap ini, hasil yang telah dihitung secara manual akan diverifikasi menggunakan aplikasi Weka, dengan tujuan untuk memastikan keselarasan antara hasil yang dicari dan hasil dari uji coba.

3.2. Objek Penelitian

Penelitian ini dilakukan di PT Excellindo Perkasa, yang bergerak dibidang penjualan komputer dan aksesorisnya. PT Excellindo Perkasa berlokasi di Komp. Bumi Indah, Lubuk Baja Kota, Kec. Lubuk Baja, Kota Batam, Kepulauan Riau 29444.



Gambar 3. 1 Lokasi Penelitian

3.3. Populasi dan Sampel

3.3.1. Populasi

Dalam konteks data mining, "populasi" mengacu pada kumpulan lengkap dari entitas atau objek yang sedang diteliti. Populasi ini dapat merujuk pada berbagai hal, tergantung pada sifat dari analisis yang dilakukan.

Dalam sebuah dataset, populasi adalah seluruh kumpulan data yang relevan yang ingin diteliti atau dianalisis. Misalnya, jika Anda sedang mengumpulkan data dari konsumen yang melakukan pembelian di sebuah toko daring, populasi dalam hal ini akan mencakup semua transaksi yang dilakukan oleh semua konsumen di toko tersebut. Adapun populasi yang digunakan peneliti adalah data transaksi penjualan di PT Excellindo Perkasa terhitung dari Oktober 2022 – Oktober 2023.

3.3.2. Sampel

Sampel dalam data mining adalah sebagian kecil dari populasi yang dipilih untuk dianalisis. Sebagian besar waktu, karena keterbatasan waktu, sumber daya, dan biaya, sulit atau tidak mungkin untuk menganalisis seluruh populasi. Oleh

karena itu, para peneliti menggunakan sampel sebagai representasi dari keseluruhan populasi.

Dalam penelitian yang berjudul "Analisis Pola Pembelian Konsumen Berdasarkan Data Historis dengan Metode Asosiasi di PT Excellindo Perkasa," sampel mungkin digunakan untuk menganalisis pola pembelian konsumen. Ini dapat dilakukan dengan cara mengambil sebagian kecil dari data historis transaksi pembelian yang dimiliki oleh PT Excellindo Perkasa.

Proses penggunaan sampel dalam penelitian ini dapat melibatkan langkah-langkah berikut:

1. Pemilihan Sampel: Dalam penelitian ini, peneliti mungkin akan memilih sampel dari data historis transaksi pembelian konsumen.
2. Analisis Data: Setelah sampel dipilih, metode asosiasi seperti Algoritma Asosiasi atau Apriori bisa diterapkan. Ini digunakan untuk menemukan pola, asosiasi, atau korelasi di antara item-item yang dibeli bersamaan. Misalnya, menemukan apakah ada kecenderungan pembelian tertentu yang sering terjadi secara bersamaan oleh konsumen.
3. Pengambilan Kesimpulan: Dari hasil analisis, peneliti dapat membuat kesimpulan tentang perilaku pembelian konsumen. Ini bisa mencakup rekomendasi strategi pemasaran, penempatan produk, atau strategi penjualan yang didasarkan pada pola pembelian yang ditemukan dari sampel tersebut.

3.4. Teknik Pengumpulan Data

Dalam penelitian "Analisis Pola Pembelian Konsumen Berdasarkan Data Historis dengan Metode Asosiasi di PT Excellindo Perkasa," teknik pengumpulan data yang tepat dan representatif menjadi kunci dalam mendapatkan hasil analisis yang akurat dan bermanfaat. Berikut adalah beberapa teknik pengumpulan data yang digunakan dalam konteks penelitian tersebut:

- Pengumpulan Data Transaksi

1. Data Historis Transaksi:

Pengumpulan data transaksi yang terdiri dari informasi pembelian konsumen dari PT Excellindo Perkasa. Ini termasuk informasi seperti ID transaksi, detail produk yang dibeli, waktu pembelian, jumlah, dan profil konsumen jika tersedia.

- Sumber Data

1. Data Internal Perusahaan:

Data historis yang dikumpulkan dari sistem informasi penjualan internal PT Excellindo Perkasa. Ini mencakup semua transaksi yang terjadi selama periode waktu tertentu.

2. Data Konsumen:

Informasi tentang preferensi konsumen, kebiasaan pembelian, atau karakteristik konsumen lainnya jika tersedia dan relevan untuk analisis.

- Teknik-Teknik Pengumpulan Data Tambahan:

1. Survei atau Wawancara:

Untuk memperoleh wawasan tambahan, mungkin dilakukan survei atau wawancara dengan sebagian konsumen untuk memahami alasan di balik keputusan pembelian atau preferensi tertentu.

2. Data Eksternal:

Memperoleh data eksternal seperti tren industri, data pasar, atau data konsumen dari sumber eksternal yang dapat memberikan konteks tambahan terhadap analisis yang dilakukan.

- **Proses Pengolahan Data:**

1. Data Cleaning:

Tahap membersihkan data dari anomali, duplikat, atau kesalahan input untuk memastikan kualitas data yang baik sebelum analisis dilakukan.

2. Transformasi Data:

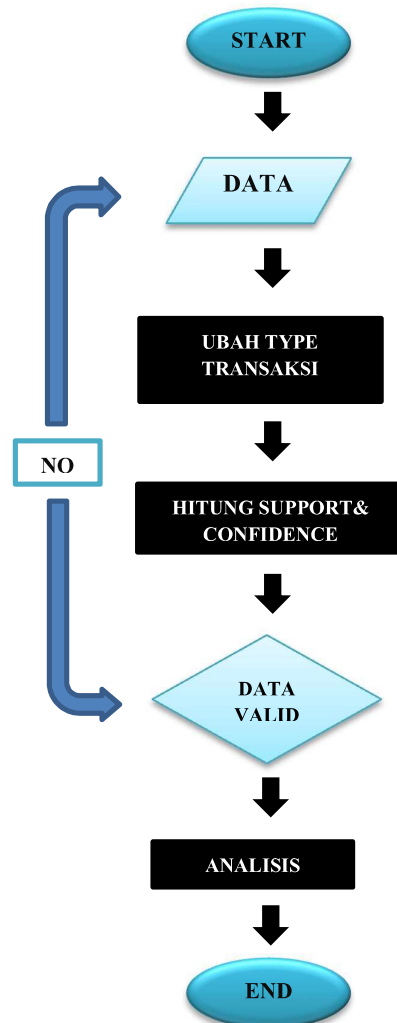
Mungkin diperlukan transformasi data untuk mempersiapkannya sesuai dengan format yang diperlukan oleh algoritma asosiasi yang digunakan. Contohnya, pembuatan itemset dari data transaksi.

3. Integrasi Data:

Mengintegrasikan data dari berbagai sumber (jika diperlukan) untuk memperoleh gambaran yang lebih lengkap.

3.5. Model Penelitian

Pengkajian model penelitian bisa ditinjau melalui flowchart dibawah ini :



(Sumber : (Riszky & Sadikin, 2019))

Dari diagram alir yang disajikan, proses pengkajian ini dimulai dengan pengumpulan data dari objek yang sedang dianalisis. Selanjutnya, data tersebut akan mengalami transformasi berdasarkan jenis data yang sesuai untuk dilakukan penambangan data. Data hasil transformasi akan diuji untuk menentukan nilai support dan confidence, serta dipilah untuk menentukan data yang memenuhi

parameter yang telah ditetapkan. Selanjutnya, pada tahap analisis, dilakukan pencarian aturan asosiasi akhir atau final.