

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 *Knowledge Discovery in Database (KDD)***

Komputer merupakan teknologi yang fungsi utamanya adalah mengelolah data untuk menghasilkan informasi. Pemrosesan data secara komputerisasi sudah diimplementasikan pada banyak bidang pekerjaan seperti pada bidang pendidikan, pemerintahan, ataupun bisnis. Data yang akan diproses ini biasanya tersimpan dalam sebuah bank data atau bisa juga kita sebut sebagai database. Tapi, dengan penggunaan database secara berkepanjangan data yang ada semakin menumpuk sementara data yang tersimpan tidak digunakan secara maksimal. Hal seperti inilah yang dapat kita sebut sebagai istilah *rich of data but poor of information*. Sebuah solusi untuk masalah ini adalah dengan menggunakan *knowledge discovery in database (KDD)*(Triantono & Suryadi, 2021).

#### **2.2 *Analysis Association***

Analisis asosiasi atau *association rule* mining adalah teknik data mining untuk menemukan aturan asosiasi antara suatu kombinasi dalam sebuah database(Desi & Agus,2022). Contoh aturan asosiasi dari analisis pembelian di suatu pasar swalayan adalah dapat diketahuinya berapa besar kemungkinan seorang pelanggan membeli roti bersamaan dengan susu. Dengan pengetahuan tersebut, pemilik pasar swalayan dapat mengatur penempatan barangnya atau merancang kampanye pemasaran

dengan memakai kupon diskon untuk kombinasi barang tertentu. Analisis asosiasi menjadi terkenal karena aplikasinya untuk menganalisis isi keranjang belanja di pasar swalayan. Ada dua ukuran parameter dalam analisa asosiasi yaitu *Support* dan *Confidence* (Nurjayadi & Kristiana, 2019).

### 1 *Support*

*Support* adalah nilai penunjang yang menunjukkan tingkat dominasi itemset dari keseluruhan transaksi, persamaanya adalah sebagai berikut

$$Support = \frac{(XUY)Count}{n} \quad \textbf{Rumus 2.1 : Rumus Support}$$

Dimana (XUY)count (dibaca x union y) adalah jumlah transaksi yang mengandung x dan y dan n adalah total transaksi yang terjadi

### 2 *Confidence*

*Confidence* adalah nilai ukuran seberapa besar valid tidaknya suatu *assosiation rule*, Persamaanya sebagai berikut

$$Confidence = \frac{(X|Y)Count}{xcount} \quad \textbf{Rumus 2.2 : Rumus Confidence}$$

Dimana Xcount adalah jumlah kemunculan item X pada seluruh transaksi

## 2.3 *Frequent Pattern Growth (FP-Growth)*

*Frequent Pattern Growth (FP-Growth)* adalah algoritma yang digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*)

dalam sebuah kumpulan data. Karakteristik algoritma *FP-Growth* adalah struktur data yang digunakan adalah tree yang disebut dengan *FP-Tree*. Dengan menggunakan *FP-Tree*, algoritma *FP-Growth* dapat langsung mengekstrak *frequent Itemset* dari *FP-Tree* (Wahyu, 2021). Penggalan itemset yang frequent dengan menggunakan algoritma *FP-Growth* akan dilakukan dengan cara membangkitkan struktur data tree atau disebut dengan *FP-Tree*

Karakteristik algoritma *FP-Growth* adalah struktur data yang digunakan adalah tree yang disebut dengan *FP-Tree*. Dengan menggunakan *FP-Tree*, algoritma *FP-Growth* dapat langsung mengekstrak frequent itemset dari *FP-Tree*. Penggalan frequent itemset dengan menggunakan algoritma *FP-Growth* dilakukan dengan cara membangkitkan struktur data tree atau disebut dengan *FP-Tree*. Metode *FP-Growth* dapat dibagi menjadi 3 tahapan utama yaitu tahap pembangkitan conditional pattern base, tahap pembangkitan conditional *FP-Tree*, dan tahap pencarian *frequent itemset*.

## **2.4 Market Basket Analysis**

*Market Basket Analysis* adalah teknik analisis yang dilakukan terhadap data penjualan untuk mengetahui produk-produk apa saja yang dibeli oleh konsumen secara bersamaan. Teknik ini merupakan salah satu bentuk aplikasi dari metode *assosiation rule*.

MBA adalah suatu metodologi yang digunakan untuk menganalisis kebiasaan belanja (*buying habit*) konsumen dengan menemukan asosiasi antar beberapa jenis barang yang berbeda, diletakkan dalam sebuah shopping basket yang

dibeli pada suatu transaksi tertentu. Tujuan dari MBA adalah untuk mengetahui produk-produk mana yang mungkin akan dibeli secara bersamaan. Analisis data transaksi dapat menghasilkan pola pembelian produk yang sering terjadi (Wulandari & Sampouw, 2020). Teknik ini telah banyak digunakan oleh toko grosir maupun retail.

MBA memanfaatkan data transaksi penjualan untuk dianalisis kemudian menemukan pola item-item yang secara bersamaan dalam suatu transaksi. Salah satu manfaat dari MBA adalah untuk merancang strategi penjualan atau pemasaran dengan memanfaatkan data penjualan yang ada di perusahaan, yaitu dengan mengubah tata letak toko, menempatkan jenis barang secara berdekatan yang sering dibeli secara bersamaan oleh konsumen dan memberikan diskon kepada barang yang jarang dibeli dan mahal. Dan dalam penelitian ini, penulis ingin memanfaatkan teknik ini untuk merancang beberapa kombinasi barang yang memiliki korelasi dalam sebuah set belanja murah.

## **2.5 SPSS**

SPSS (*Statistical program for Social Science*) merupakan program aplikasi yang dapat digunakan untuk melakukan pengolahan data statistik. Aplikasi ini dapat digunakan untuk hampir seluruh jenis file data dan dapat sekaligus membuat laporan dalam bentuk tabulasi, grafik, dan plot untuk berbagai distribusi maupun statistik deskriptif (Laksamana et al., 2021).

## **2.6 Aplikasi WEKA**

*Waikato Environment for Knowledge Analysis (WEKA)* merupakan rangkaian perangkat lunak pembelajaran mesin yang ditulis dalam bahasa java yang

digunakan untuk keperluan penambangan data. Weka juga memiliki banyak tool untuk pengolahan data mulai dari *pre processing*, *classification*, *association rules*, dan *visualization*(Faid et al., 2019).

## 2.7 Penelitian Terdahulu

Penelitian data mining ini tentunya menggunakan referensi dari beberapa penelitian yang telah dilakukan oleh peneliti sebelumnya. Adapun referensi jurnal sebagai berikut:

1. Penelitian oleh (Suhada et al., 2020) dari Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri, E-ISSN: 2549-5178 Vol.8 No.2, dengan judul **“Penerapan Algoritma FP-Growth untuk Menentukan Pola Pembelian Konsumen Pada AHASS Cibadak”**. Penelitian ini bertujuan untuk menentukan pola pembelian konsumen dengan salah satu algoritma asosiasi yaitu *FP-Growth* dengan data transaksi penjualan di PT. Selamat Lestari Mandiri Cibadak. Data transaksi penjualan memiliki 13 atribut dan 216 catatan. Berdasarkan penelitian yang diperoleh dari hasil penjualan suku cadang, terdapat beberapa item produk yang paling banyak dijual secara simultan serta bersamaan di PT. Selamat Lestari mandiri Cibadak. Dengan algoritma *FP-Growth* untuk menganalisa pola pembelian konsumen sangat bermanfaat bagi perusahaan, karena PT. Selamat Lestari Mandiri Cibadak akan mengetahui suku cadang mana yang banyak dibeli secara simultan serta bersamaan dan membantu dalam pemesanan suku cadang pada kantor pusat.

2. Penelitian oleh (Putri & Mandala, 2021) dari Universitas Putera Indonesia YPTK Padang, ISSN: 2548 – 8368, Vol.5 No.1, dengan judul **“Implementasi Algoritma FP-Growth Untuk Menemukan Pola Frekuensi Pembelian Lauk Pada Rumah Makan Takano Joo”**. Penelitian ini menggunakan data transaksi penjualan lauk dalam satu hari sebagai data sampel sebanyak 12 transaksi. Contoh dari pola beli yang dihasilkan adalah dari 12 transaksi yang membeli ayam bakar, 6 transaksi membeli rendang daging. Algoritma yang digunakan untuk membantu rumah makan dalam menemukan pola beli pelanggan adalah Algoritma *FP-Growth*, sehingga dapat membantu rumah makan dalam mengatur jumlah lauk yang akan dijual setiap harinya. Hasil dari penelitian ini adalah keputusan dalam pengadaan lauk di rumah makan yang dapat dilihat dari pola frekuensi pembelian lauk yang dilakukan oleh pelanggan.
3. Penelitian oleh (Wijaya, 2020) dari Fakultas Ilmu Komputer Unsri, E-ISSN 2503 -2933, Vol.7 No.2, dengan judul **“Analisa Pola Frekuensi Keranjang Belanja Dengan Perbandingan Algoritma FP-Growth (Frequent Pattern Growth) Dan Eclat pada Minimarket”**. Penelitian ini melakukan analisis pola data transaksi penjualan dengan menerapkan metode asosiasi pada data mining. Data dipersiapkan dengan melalui proses seleksi data, pembersihan data dan ditransformasi ke bentuk yang dapat diolah oleh sistem. Selanjutnya data diolah menggunakan algoritma *FP-Growth* dan *Eclat* dengan

melakukan perbandingan hasil dari algoritma berupa aturan asosiasi yang terbentuk dan kecepatan waktu proses dari masing-masing algoritma dengan minimum *support* dan *confidence* sebesar 0.01% untuk menentukan jumlah aturan yang kuat pada tabel dan grafik sebagai bahan pengambil keputusan yang ditunjukkan untuk frekuensi keranjang belanja.

4. Penelitian oleh (Haya & Ramadani, 2020) dari Program Studi Teknik Informatika Universitas Samudra dengan judul **“Penerapan Data Mining dalam Analisis Pola Pembelian Minuman dan Makanan di Kantin SMAN 4 Langsa Menggunakan Algoritma FP-Growth”**. Hasil penelitian ini memiliki dua bentuk, aturan yang dihitung secara manual dan aturan melalui Penambang cepat. Kemudian hasilnya akan digunakan untuk pengelola kantin untuk mempelajari pola siswa dalam membeli minuman dan makanan di kantin.
5. Penelitian oleh (Krisma, 2020) dari Universitas Teknokrat Indonesia dengan judul **“ Penerapan Data Mining Untuk Analisis Pola Pembelian Pelanggan Menggunakan Algoritma Apriori (Studi Kasus: Toko Diengva Bandar Jaya)”**. Menerapkan Algoritma Apriori sebagai metode untuk mencari pola pola beli pelanggan pada data transaksi penjualan Toko Diengva menggunakan *Rapid Miner*. Dengan aturan yang dihasilkan dari penerapan Algoritma Apriori dapat dijadikan sebagai dasar untuk melakukan stok terhadap item-item yang memenuhi nilai *minimum support* dan *minimum confidence*. Item-item

yang memenuhi aturan tersebut adalah Bulu Mata, Lemb Bulu Mata, Soflen, dan Air Soflen. Nilai *confidence* dari hubungan antar 2 item diatas dapat dikatakan tinggi sehingga hasil aturan tersebut dapat dijadikan sebagai dasar untuk melakukan stok terhadap item-item diatas.

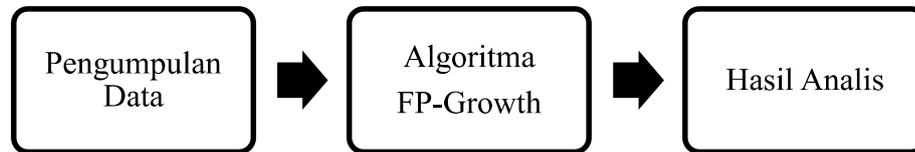
6. Penelitian oleh (Mustika et al., 2021) dari Universitas Islam Negeri Sultan Syarif Kasim Riau dengan judul **“Analisis Keranjang Pasar Menggunakan Algoritma K-Means dan FP-Growth pada PT. Citra Mustika Pandawa”**. Hasil dari penelitian ini adalah pola pembelian konsumen pada PT. Citra Mustika Pandawa ditarik kesimpulan setelah dilakukan beberapa percobaan, dan dari percobaan itu diperoleh cluster optimal sebanyak 5 Cluster berdasarkan nilai uji validitas Davies Bouldin Indeks (DBI) sebesar 0.500. Ditemukan juga bahwa terdapat 8 items/produk yang sering dibeli oleh panggan pada PT. Citra Mustika Pandawa yaitu produk Meja Makan *Elite*, Rak Piring Asia Jaya, TV Led 32”, Lemari Es Ip, *Kitchen Set*, Mesin Cuci 2T, dan Lemari Pakaian IP.
7. Penelitian oleh ( Desi & Agus,2022) dengan judul **“Implementasi Market Basket Analysis Menggunakan Association Rule Menerapkan Algoritma FP-Growth”**. Hasil penelitian ini adalah penggunaan metode asosiasi *fp-growth* terbukti mampu mengetahui hasil analisis yang muncul secara bersamaan berdasarkan data penjualan obat di Apotek French Farma dengan hasil perhitungan

keduanya yang terpilih adalah 5 item barang yang memenuhi minimum *support* 20 % dan minimum *confidence* 90 %.

8. Penelitian oleh (Harman, 2020) dari Universitas Putera Batam dengan judul **“Penerapan Data Mining Untuk Prediksi Penjualan Ready Mix Menggunakan Metode Algoritma C4.5 Pada PT Remicon Widyaprima”**. Hasil penelitian ini, data yang diperoleh dapat bekerja secara akurat pada aturan setiap tes, ini dapat memudahkan penelitian berbasis pada data yang telah diperoleh melalui proses pengujian prediksi penjualan yang dikategorikan dalam beberapa variabel. Pengiriman barang, kualitas produk, harga produk, jaminan / ketahanan beton readymix, tipe readymix dan kapasitas. Peneliti menggunakan 2 variabel keputusan yaitu membeli readymix dan tidak membeli readymix.

## **2.8 Kerangka Pemikiran**

Kerangka Pemikiran merupakan suatu diagram yang menjelaskan rencana atau alur logika untuk berjalannya sebuah penelitian yang diangkat. Konsep di dalam kerangka pemikiran itu juga menjelaskan mengenai suatu hubungan antara variabel yang satu dengan variabel yang lainnya. Melalui pemanfaatan data mining menggunakan metode algoritma *FP-Growth* di toko Lovebaby Store dapat ditentukan susunannya, tahapan yang dibutuhkan untuk mendapatkan hasil rancangan dengan fungsi yang baik sebagai berikut:



**Gambar 2. 4** Kerangka Pemikiran

**Sumber:** (Data Peneliti, 2022)

Seperti yang terlihat dari *flowchart* diatas, tahapan awal pada penelitian ini yaitu:

1. Dihimpun lalu dilaksanakan cleaning data, transformation guna mengetahui pola pengambilan data. Tahapan tersebut adalah pengambilan data yang merupakan transaksi pembelian di toko Love Baby Store.
2. Tahap kedua ialah proses analisis data dengan *FP-Growth* dengan tahap berikut:



**Gambar 2.5** Tahapan Analisis Data

**Sumber:** (Data Peneliti, 2022)

3. Temuan penelitian ini adalah informasi mengenai pola pembelian oleh konsumen kemudian di susun dalam paket-paket item belanja yang relatif lebih murah.