

BAB 2

KAJIAN PUSTAKA

1.1 Penelitian Terkait

Pada penelitian ini, kajian pustaka yang dilakukan untuk mendukung ide penelitian. Penelitian yang telah dilakukan sebelumnya yaitu :

2.1.1 Penelitian pertama “Memaksimalkan Potensi Penjualan Ikan Menggunakan *Data Mining*” Oleh (Ilham Yori Pradana dkk). Pada penelitian tersebut menemukan permasalahan bagaimana cara memaksimalkan penjualan ikan. Metode yang digunakan pada penelitian tersebut yaitu *data mining* dan *market basket analysis*. Dan untuk algoritma apriori digunakan untuk membantu menemukan sejumlah aturan asosiasi dari basis data transaksi penjualan ikan.

Hasil dari penelitian tersebut yaitu menghasilkan pola atau *itemset* dari data *dummy* yang telah digunakan dari metode *market basket analysis* ditemukan dua pola, pola yang pertama jam alarm yang berwarna merah dan hijau, sedangkan pola kedua yaitu, gelas kertas berwarna merah, serbet dan piring. Akan tetapi keberhasilan dari penelitian ini masih belum bisa ditentukan karena, belum diterapkannya metode *hottelling* dan juga perhitungan peningkatan pembelian masih belum diketahui.

2.1.2 “*Market Basket Analysis* Pada Mini Market Ayu Dengan Algoritma Apriori” Oleh (Sistem & Elisa, 2018). Masalah yang ditemukan pada penelitian tersebut adalah data transaksi penjualan pada mini market Ayu hanya digunakan sebagai arsip tanpa dimanfaatkan dengan baik. Metode yang digunakan yaitu Analisis keranjang pasar dengan Algoritma Apriori.

Hasil penelitian menemukan bahwa nilai dukungan dan kepercayaan tertinggi adalah Minyak dan Susu dengan nilai dukungan 42,85% dan kepercayaan 85,71%.

2.1.3 “Analisis Startegi Pemasaran Menggunakan Metode *Assosiation Rule – Market Basket Analysis* Untuk Menentukan Komoditas Utama Penjualan Serta Mengoptimalkan Sumber Daya Agar Dapat Memenuhi Permintaan Pasar” Oleh (Pramuditya & , Mahendra Aji Wicaksana, 2018). Pada penelitian tersebut ditemukan permasalahan yang ditemui yaitu Fanri Collection belum mampu memetakan jenis produk yang seharusnya dioptimalkan dalam pembuatannya sehingga produk tersebut dapat mudah terjual dan mengurangi *inventory* produk-produk yang kurang laku di pasaran. Metode yang digunakan yaitu *assosiation rule-market basket analysis*.

Hasil *Rule* kombinasi yang didapatkan dengan minimal *support* sebesar 0,5 dan minimal *confidence* sebesar 0,5 , ada 3 *rule* yang dianggap valid karena nilai *Lift ratio* > 1. Jenis produk tas wanita akan terbeli bersamaan dengan jenis produk sabuk dengan tingkat kepercayaan 60% dan didukung sebesar 12% dari nilai keseluruhan. Jenis produk dompet wanita tipe 2 akan terbeli bersamaan dengan jenis produk dompet wanita tipe 3 tingkat kepercayaan 76% dan didukung sebesar 20% nilai keseluruhan. Jenis produk gantungan kunci akan terbeli bersamaan dengan jenis produk dompet wanita tipe 3 dengan tingkat kepercayaan 85,7% dan didukung sebesar 12% dari nilai keseluruhan.

2.1.4 “Implementasi *Data Mining* Pola penjualan Sparepart Motor Honda Pada Pt Rotella Persada Mandiri Dengan Menggunakan Algoritma Apriori” Oleh (Elyas & Prayoga, 2020). Suku cadang atau *sparepart* yaitu suatu alat yang mendukung pengadaan barang untuk keperluan peralatan yang digunakan dalam proses produksi”. Berdasarkan definisi diatas, suku cadang merupakan faktor utama yang menentukan jalannya proses produksi dalam suatu perusahaan. Metode yang digunakan pada penelitian tersebut yaitu *data mining* dan *market basket analysis*. Dan untuk algoritma apriori digunakan untuk membantu menemukan sejumlah aturan asosiasi.

Hasil dari penelitian ini yaitu bahwa dengan memanfaatkan algoritma apriori untuk menghasilkan sebuah data baru yang dapat digunakan sebagai referensi bagi pemilik toko barang dalam hal penyajian barang kepada konsumen

sebagai bagian dari memaksimalkan penjualan.

2.1.5 “Penerapan *Data Mining* dengan Memanfaatkan Metode *Association Rule* untuk Promosi Produk” Oleh (Rerung, 2018). Penelitian ini bertujuan untuk menerapkan *data mining* untuk promosi produk Distro Nasional. Dalam bidang keilmuan *data mining*, terdapat suatu metode yang dinamakan *association rule*. Metode ini bertujuan untuk menunjukkan nilai asosiatif antara jenis-jenis produk yang dibeli oleh pelanggan sehingga terlihatlah suatu pola berupa produk apa saja yang sering dibeli oleh pelanggan tersebut. Dengan mengetahui jenis produk yang sering dibeli maka dapat dibuat sebagai sebuah dasar keputusan untuk menentukan produk apa saja yang cocok untuk dipromosikan kepada pelanggan tersebut. Algoritma Apriori juga akan dipergunakan untuk menentukan *frequent item set* sehingga hasil akhir yang dicapai yaitu untuk menghitung persentase *confidence* pelanggan terhadap produk yang ditawarkan.

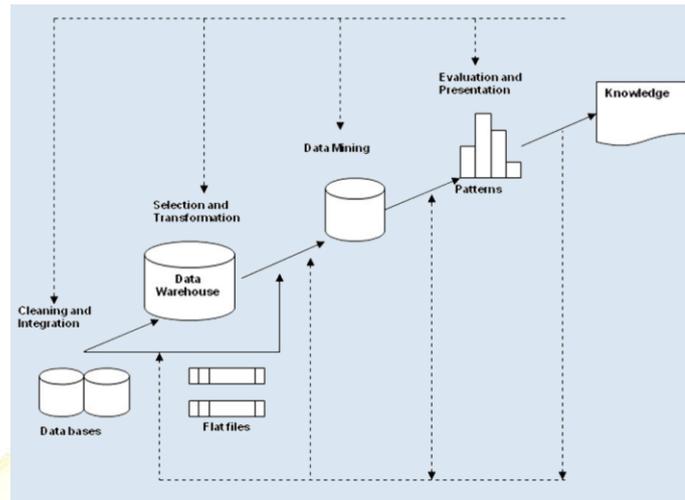
2.2 Landasan Teori

Dalam sebuah penelitian terdapat landasan dalam penalaran teori yang digunakan, adapun landasan teori pada penelitian ini :

2.2.1 Data Mining

Data Mining ialah sebuah proses yang menggunakan teknik *statistic*, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstrasi dan mengidentifikasi informasi yang bermanfaat serta pengetahuan yang terkait dari berbagai *database* besar. *Data mining* juga sering disebut sebagai *Knowledge Discovery in Database* (KDD), yang berarti sebuah kegiatan yang meliputi pengumpulan dan pemakaian data historis untuk menemukan keteraturan, pola atau hubungan antar data dalam set data berukuran 24 besar. Keluaran *data mining* ini bisa dipakai untuk memperbaiki pengambilan keputusan di masa depan (Umayah, 2015).

Proses yang harus dilalui untuk mendapatkan suatu informasi dapat digambarkan pada **Gambar 2.1**.



Gambar 2. 1 Proses *data mining*

1. *Data Cleaning*. Proses memperbaiki atau menghapus data, set, tabel, dan *database* yang tidak akurat.
2. *Data Integration*. Proses penggabungan data yang telah diseleksi yang nantinya akan disimpan kedalam *database*.
3. *Data Selection*. Penyeleksian data yang relevan dengan penelitian yang diambil dari *database*.
4. *Data Transformation*. Proses transformasi dan konsolidasi data ke dalam berbagai bentuk yang sesuai untuk data mining.
5. *Data Mining*. Proses mengekstraksi data untuk menemukan sebuah pola, menggunakan metode-metode tertentu.
6. *Pattern Evaluation*. Mengidentifikasi pola yang menarik dan yang mewakili informasi dari data.
7. *Knowledge*. Informasi yang telah jadi yang nantinya akan ditampilkan kepada pengguna dengan cara yang menarik (Pradana, n.d.)

2.2.2 Pengelompokan *Data Mining*

Data Mining dikelompokkan dalam beberapa kategori berdasarkan fungsi, yaitu *Description*, *Estimation*, *Prediction*, *Classification*, *Clustering*, dan *Association*.

1. Deskripsi (*Description*)

Merupakan suatu fungsi yang bertujuan memahami lebih lanjut mengenai data yang diamati sehingga dapat diketahui perilaku dari sebuah data.

2. Estimasi (*Estimation*)

Estimasi adalah suatu pengukuran yang didasarkan pada hasil kuantitatif atau hasil yang akurasi dapat diukur dengan angka.

3. Prediksi (*Prediction*)

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

4. Klasifikasi (*Classification*)

Bertujuan untuk menyimpulkan beberapa definisi karakteristik dari sebuah grup. Misalnya, pelanggan perusahaan yang sudah berpindah karena tersaingin oleh perusahaan lain.

5. Pengklasteran (*Clustering*)

Clustering adalah identifikasi kelompok dari produk-produk atau barang-barang yang memiliki karakteristik khusus.

6. *Association*

Association merupakan identifikasi hubungan dari kejadian-kejadian yang sudah terjadi di suatu waktu.

2.2.3 Market Basket Analysis (MBA)

Market Basket Analysis merupakan salah satu metode yang biasa digunakan dan paling bermanfaat pada lingkungan marketing untuk keperluan strategi pemasaran, desain katalog serta proses pembuatan keputusan bisnis.

Salah satu sumber data dari market basket analysis dapat diperoleh dari sebuah transaksi yang telah dilakukan. *Market basket analysis* umumnya dimanfaatkan sebagai titik awal pencarian pengetahuan dari suatu

data transaksi ketika belum diketahui pola spesifik yang akan dicari. Kebutuhan *market basket analysis* berawal dari keakuratan dan manfaat yang dihasilkannya dalam bentuk suatu aturan assosiasi (Umayah, 2015).

Market basket analysis memiliki suatu konsekuensi yang perlu diperhatikan dalam penggunaannya, yaitu semakin banyak produk barang yang terdapat di *database* akan menyebabkan *rules* yang dihasilkan oleh *market basket analysis* semakin meningkat. Dari sekian banyak *rules* tersebut, hanya beberapa *rules* saja yang bersifat informatif dan mampu mengungkap pola transaksi yang sebelumnya tidak ada. Berdasarkan hal ini, maka diperlukan suatu proses pemilihan *rules* yang didasarkan pada suatu ukuran kebaikan. Ukuran kebaikan yang umumnya digunakan untuk pemilihan *rules* adalah *support*, *confidence*. Semakin besar kedua nilai tersebut, maka *rules* tersebut semakin baik dalam memprediksi pola pembelian konsumen. "Suatu rule dikatakan menarik dan memiliki kemampuan prediksi yang baik apabila *rule* ini memiliki nilai ukuran kebaikan yang tinggi dalam *database*.

Ukuran kebaikan yang pada umumnya digunakan adalah *support*, *confidence*. Ukuran kebaikan *support* mengukur proporsi kemunculan item dalam transaksi.

Secara *matematis*, nilai *support* sebuah item diperoleh dengan rumus:

$$\text{Support (A)} = \frac{\text{Jumlah Transaksi mengandung A}}{\text{Total transaksi}}$$

Sedangkan nilai *support* dengan dua item diperoleh dengan rumus :

$$\text{Support (A} \cap \text{B)} = \frac{\text{Jumlah Transaksi mengandung A dan B}}{\text{Total Transaksi}}$$

Dan Nilai *support* dengan tiga item diperoleh dengan rumus dibawah ini :

$$\text{Support (A} \cap \text{B} \cap \text{C)} = \frac{\text{Jumlah Transaksi mengandung A, B dan C}}{\text{Total Transaksi}}$$

Ukuran kebaikan *confidence* mengukur kemampuan prediksi atau tingkat kepercayaan dari suatu *rule*. Secara matematis *confidence* dituliskan sebagai :

$$CONFIDENCE = \frac{\sum \text{item yang dibeli sekaligus}}{\sum \text{jumlah transaksi pada bagian antecedent}}$$

Suatu *rule* dikatakan baik apabila memiliki nilai *support* dan *confidence* yang mendekati 1 (atau 100%) dan *lift* di atas nilai 1 (Pradana, n.d.).

2.2.4 Algoritma Apriori

Algoritma apriori adalah algoritma yang paling terkenal untuk menemukan pola frekuensi tinggi. algoritma apriori dibagi menjadi beberapa tahap yang disebut narasi atau pass Pembentukan kandidat itemset, kandidat *k-itemset* dibentuk dari kombinasi (*k-1*)-*itemset* yang didapat dari iterasi sebelumnya. Satu cara dari algoritma apriori adalah adanya pemangkasan kandidat *k-itemset* yang subset-nya yang berisi (*k-1*) item tidak termasuk dalam pola frekuensi tinggi dengan panjang *k-1*.

Algoritma apriori adalah salah satu algoritma yang melakukan pencarian *frequent itemset* dengan menggunakan teknik *association rule*. Untuk mencari *association rule* dari suatu kumpulan data, tahap pertama yang harus dilakukan adalah mencari *frequent itemset* terlebih dahulu. *Frequent itemset* adalah sekumpulan item yang sering muncul secara bersamaan. Penting tidaknya suatu asosiasi dapat diketahui dengan dua tolak ukur, yaitu *support* dan *confidence*. *Support* adalah nilai penunjang atau persentase kombinasi sebuah item dalam *database*, sedangkan *confidence* adalah nilai kepastian yaitu kuatnya hubungan antar item dalam aturan asosiasi. Proses utama yang dilakukan dalam algoritma apriori untuk mendapat *frequent itemset* (Sistem & Elisa, 2018).