

BAB I

PENDAHULUAN

1.1. Latar Belakang

Salah satu cara yang umum dipakai menganalisis data penjualan adalah dengan mencari *frequent itemset*. Jadi, *frequent itemset* merupakan kumpulan barang yang sering dibeli bersamaan dalam transaksi. Biasanya, digunakan untuk menemukan *itemset* yang sering muncul, dapat memakai algoritma Apriori. Algoritma ini adalah bagian dari teknik dalam aturan asosiasi dalam data mining, yang fungsinya mencari hubungan antara berbagai item dalam data transaksi[1]. Namun, dalam praktiknya tidak semua *frequent itemset* memiliki nilai yang sama bagi bisnis. Oleh karena itu, muncul kebutuhan untuk mengidentifikasi *top-K frequent itemset*, yaitu K *itemset* yang paling sering muncul dan memiliki dampak terbesar bagi bisnis.

Rumah Cantik Marwah, yang berlokasi di Jl. Hayam Wuruk Jl. Ngaljo, Sitimulyo, Kecamatan Cepu, Kabupaten Blora, Jawa Tengah, adalah sebuah toko yang menawarkan berbagai jenis produk kosmetik. Produk-produk yang dijual di toko ini mencakup merek-merek seperti Marwah, Hi Glow, Skin Relieve 12 (SR12), Ratu Arab, Henna Naura, dan lain-lain. Rumah Cantik Marwah tidak hanya berperan sebagai distributor dari produk-produk tersebut, tetapi juga menawarkan berbagai layanan perawatan kecantikan seperti *facial reguler*, *facial biolight*, dan lainnya.

Dalam operasionalnya, Rumah Cantik Marwah sudah menggunakan *software* untuk mengumpulkan data. Data yang dikumpulkan oleh *software* tersebut berupa data keseluruhan penjualan dalam bentuk *pdf*. Namun dalam proses tata kelola data, Rumah Cantik Marwah belum melakukan pengelolaan mengenai pola pembelian konsumen sehingga diperlukan analisis data penjualan. Analisis data penjualan dapat memberikan informasi yang berharga mengenai kebiasaan dan preferensi konsumen.

Penelitian terdahulu salah satunya "Penerapan Asosiasi Menggunakan Algoritma Apriori Untuk Menganalisis Pola Pembelian Konsumen Pada Toko Kelontong". Pada penelitian ini Minimarket Mawar Simp.Tangsi Balimbingan memiliki data penjualan yang belum dimanfaatkan secara optimal. Sehingga

analisis data diperlukan untuk mengidentifikasi pola penjualan dalam periode tertentu menggunakan algoritma apriori. Hasilnya, dengan minimum *support* 30% dan *confidence* 80%, terbentuk 8 aturan asosiasi[2].

Penelitian berjudul "Menentukan Akurasi Tata Letak Barang dengan Menggunakan Algoritma Apriori dan Algoritma *FP-Growth*" bertujuan untuk menemukan kesamaan antara barang berdasarkan item yang dibeli, yang bisa jadi panduan untuk tata letak dan identifikasi barang saat menambah stok. Dua metode yang digunakan adalah algoritma Apriori dan *FP-Growth*. Hasilnya, algoritma Apriori menghasilkan 4 aturan dengan *support* 1.31296 dan akurasi 78%, sementara *FP-Growth* menghasilkan 6 aturan dengan *support* 1.67938 dan akurasi 128%[3].

Selain itu, penelitian berjudul "Penerapan Data Mining Untuk Menentukan Pola Penjualan Menggunakan Metode Algoritma Apriori (Studi Kasus: Coffe Shop Kopi Bonjera Jakarta)" yang fokus pada masalah untuk menemukan pola penjualan yang tepat untuk promosi. Dalam penelitian ini, memakai data mining dengan algoritma Apriori untuk melihat hubungan antara penjualan produk berdasarkan kebiasaan dan selera konsumen. Caranya, dengan memasukkan data penjualan dari periode tertentu. Hasilnya, ditemukan beberapa aturan yang memenuhi 3 syarat penting dari algoritma Apriori, yaitu: *support* relatif lebih dari 16,67%, *confidence* lebih dari 20%, dan *lift ratio* yang sama atau lebih dari 1. Dua contohnya adalah: 1. Jika orang membeli kentang goreng *cheese*, biasanya mereka juga membeli kopi susu gula aren, 2. Jika mereka membeli leci *ice tea*, mereka cenderung juga membeli kopi susu gula aren[4].

Pada penelitian "*Multi-threaded approach in generating frequent itemset of Apriori algorithm based on trie data structure*" membahas tentang penerapan struktur data *multi-threaded* dan *trie* pada masalah perhitungan dukungan dalam algoritma Apriori. Hasil penelitian menunjukkan bahwa model *multi-threaded* yang menerapkan struktur varian *double* SPMD dapat bekerja hampir tiga kali lebih cepat daripada model *multi-threaded* yang menerapkan struktur SPMD, struktur MPMD, dan kombinasi MPMD dan SPMD berdasarkan perbedaan waktu dari hasil eksperimen 5-*itemsets* dan 10-*itemsets*[5].

Penelitian sebelumnya yang berjudul "Analisis Pola Penjualan Produk Makanan dan Minuman Menggunakan Algoritma Apriori" membahas masalah

bahwa data transaksi penjualan di Kafe Aksara tidak dimanfaatkan untuk membuat strategi penjualan yang bisa menarik pelanggan, sehingga hanya menjadi arsip yang tidak berguna. Untuk mengatasi ini, mereka melakukan analisis memakai data mining dengan algoritma Apriori. Hasilnya ada 14 aturan, dan bisa disimpulkan bahwa produk makanan dan minuman yang sering terjual bersamaan adalah *Dimsum Ayam*, *Taro*, *Fried Fries*, *Lemon Tea*, dan *Red Velvet*, dengan persentase antara 81,2% sampai 94,1%[6].

Pada penelitian "*An efficient apriori algorithm for frequent pattern mining using mapreduce in healthcare data*" memiliki tujuan penelitian untuk mengurangi permintaan sumber daya yang besar dan mengurangi komunikasi *overhead* ketika data frekuen diekstraksi. Sehingga algoritma Apriori yang ditingkatkan dengan nama *enhanced parallel and distributed apriori* (EPDA) disajikan untuk industri perawatan kesehatan, berdasarkan lingkungan yang dapat diskalakan yang dikenal sebagai *Hadoop MapReduce*. Hasil dari penelitian di mana EPDA berkinerja dalam hal waktu dan jumlah aturan yang dihasilkan dengan basis data kesehatan dan nilai dukungan minimum yang berbeda[7].

Berdasarkan penelitian terdahulu penulis mengambil judul penelitian "PENGALIAN *TOP-K FREQUENT ITEMSET* MENGGUNAKAN ALGORITMA APRIORI PADA DATASET PENJUALAN KOSMETIK" didasari oleh relevansi praktis dan efektivitas algoritma Apriori yang terbukti dalam mengidentifikasi pola penjualan signifikan, seperti yang ditunjukkan oleh penelitian-penelitian sebelumnya. Studi-studi tersebut menggarisbawahi kemampuan algoritma ini dalam membantu pengambilan keputusan bisnis, mengoptimalkan penggunaan data transaksi yang sering kali hanya menjadi arsip, dan meningkatkan efisiensi analisis data. Dengan fokus pada industri kosmetik yang memiliki karakteristik dan dinamika khusus, penelitian ini dapat memberikan informasi lebih lanjut tentang preferensi pelanggan dan mendukung strategi pemasaran yang lebih personal dan efektif, sehingga dapat memaksimalkan penjualan dan kepuasan pelanggan.

Penelitian ini memfokuskan dalam menggali *top-K frequent itemset* dari dataset penjualan kosmetik di Rumah Cantik Marwah dengan menggunakan algoritma apriori. Adanya data penjualan yang tepat, penelitian ini memberikan

informasi relevan tentang pola belanja konsumen kosmetik dan memberi rekomendasi yang dapat membantu perusahaan agar merancang strategi pemasaran yang lebih tepat. Misal, *bundling* produk, menempatkan produk di tempat yang strategis, dan melakukan promosi yang lebih pas sasaran. Selain itu penggalian *top-K frequent itemset* dapat membantu perusahaan untuk lebih fokus pada *itemset* yang benar-benar penting dan relevan.

1.2. Perumusan Masalah

1.2.1. Permasalahan

Rumah Cantik Marwah, sebagai distributor dan penyedia layanan perawatan kecantikan, dalam operasinya sudah menggunakan *software* untuk mengumpulkan data. Data yang dikumpulkan oleh *software* tersebut berupa data keseluruhan penjualan dalam bentuk *pdf*. Namun, pemanfaatan data tersebut masih belum optimal dalam memberikan informasi yang mendalam tentang pola pembelian konsumen. Sehingga permasalahan yang dihadapi adalah belum adanya identifikasi yang jelas mengenai item-item yang sering dibeli bersamaan oleh konsumen, serta perlunya metode untuk mengidentifikasi *itemset* yang paling sering muncul dan memiliki dampak terbesar bagi bisnis.

1.2.2. Metode Usulan

Metode yang diusulkan untuk mengatasi masalah ini adalah menggunakan algoritma Apriori untuk mencari *top-K frequent itemset* dari data penjualan kosmetik di Rumah Cantik Marwah. Algoritma apriori merupakan salah satu cara dalam data mining yang digunakan mencari *association rules*, yang berarti menunjukkan hubungan antara berbagai atribut.

1.2.3. Pertanyaan penelitian

Dari penjelasan di atas, penelitian ini akan berfokus ke beberapa pertanyaan penelitian berikut ini:

1. Bagaimana pola *association rules* pembelian konsumen di Rumah Cantik Marwah berdasarkan data penjualan yang ada?
2. Bagaimana algoritma Apriori dapat diterapkan untuk mengidentifikasi *top-K frequent itemset* dalam *dataset* penjualan kosmetik?

1.3. Tujuan dan Manfaat

1.3.1. Tujuan Penelitian

Dari masalah yang sudah dibahas, ada beberapa tujuan di penelitian ini, yaitu:

1. Mengidentifikasi pola pembelian konsumen di Rumah Cantik Marwah berdasarkan data penjualan yang ada.
2. Menerapkan algoritma Apriori untuk mengidentifikasi *top-K frequent itemset* dalam dataset penjualan kosmetik.

1.3.2. Manfaat Penelitian

Manfaat dalam penelitian ini antara lain:

1. Untuk Rumah Cantik Marwah, hasil penelitian ini bisa digunakan untuk strategi pemasaran dengan menganalisis pola pembelian konsumen dan *itemset* yang sering dibeli bersamaan. Hal tersebut meningkatkan penjualan dan membantu dalam pengambilan keputusan yang lebih baik.
2. Peneliti mendapatkan pengalaman langsung dalam mengolah dan menganalisis data penjualan dengan menggunakan algoritma Apriori, yang merupakan teknik penting dalam data mining.
3. Bagi peneliti selanjutnya, hasil penelitian ini merupakan bahan untuk mengembangkan penelitian di bidang data mining lebih lanjut.

1.4. Batasan Masalah

Batasan masalah pada penelitian ini antara lain:

1. Penelitian ini hanya memakai dataset penjualan dari satu perusahaan yaitu Rumah Cantik Marwah Cepu.
2. Menggunakan metode Algoritma Apriori.
3. Dataset transaksi penjualan bulan Januari-Maret 2024 adalah data yang digunakan dalam penelitian ini.
4. Hasil pada penelitian ini berupa *top-K itemset* yang dapat direkomendasikan untuk perusahaan.

1.5. Sistematika Penulisan

Terdapat beberapa bab dalam penulisan penelitian ini, yaitu:

BAB I PENDAHULUAN

Bab ini menguraikan latar belakang penelitian yang menjelaskan masalah yang dihadapi, tujuan yang ingin dicapai, manfaat dari penelitian, perumusan masalah yang menjadi fokus penelitian, serta gambaran umum mengenai penulisan laporan.

BAB II KAJIAN PUSTAKA

Bab ini membahas berbagai konsep, teori, dan definisi yang mendasari penelitian. Kajian pustaka ini mencakup literatur yang relevan sebagai landasan teoritis untuk memahami dan mendukung analisis dalam penelitian.

BAB III METODE USULAN

Bab ini menguraikan metode penelitian yang digunakan, mencakup lokasi dan waktu penelitian, serta tahapan-tahapan dalam penerapan metode.

BAB IV HASIL DAN PEMBAHASAN

Di bab ini, hasil penelitian dipaparkan dan dianalisis. Pembahasan dilakukan untuk menginterpretasikan hasil yang diperoleh sesuai dengan tujuan dan rumusan masalah penelitian, serta mengaitkannya dengan teori yang relevan.

BAB V PENUTUP

Bab ini berisi kesimpulan yang diambil dari hasil penelitian serta saran yang dapat diberikan untuk penelitian lebih lanjut atau penerapan praktis dari hasil penelitian.

REFERENSI

Bagian ini mencantumkan semua sumber pustaka dan referensi yang digunakan dalam penelitian sebagai dasar teoretis dan pendukung argumen.