

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Meskipun teknologi informasi khususnya komputasi telah diadopsi di berbagai bidang, namun masih banyak perusahaan yang belum bisa memanfaatkannya secara optimal. Perusahaan biasanya memiliki data historis transaksi penjualan untuk setiap bulannya, namun sayangnya hanya digunakan sebagai laporan mingguan dan bulanan. Sebagian besar perusahaan juga menggunakan cara manual dalam melakukan strategi pemasaran produknya, cara-cara tersebut tidak memiliki acuan dan hanya berdasarkan perkiraan. Mekanisme seperti ini dapat menyebabkan promosi yang tidak tepat dan pemborosan biaya promosi. Data transaksi penjualan yang tersimpan dapat membawa keuntungan bagi manajemen perusahaan, seperti meningkatkan penjualan melalui strategi promosi produk yang tepat. Sebagian besar konsumen tertarik dengan program promosi karena mereka merasa dapat berbelanja dengan lebih efektif. Kondisi tersebut dapat menjadi sarana untuk meningkatkan penjualan, mempertahankan pelanggan, dan memuaskan pelanggan.

Proses pencarian informasi penting yang tersembunyi di dalam data yang besar dengan cara menggali pola informasi menggunakan metode tertentu, yang dimana informasi tersebut digunakan untuk hal yang berguna dan mempermudah dalam pengambilan keputusan disebut data mining [1]. Dengan melakukan penggalian pola informasi terhadap data yang berguna untuk mengubah data tersebut menjadi sebuah informasi yang baru dan bermanfaat yang didapatkan melalui cara mengeksplorasi juga mencari nilai-nilai yang berharga yang didapat dari data didalam basis data. *Frequent pattern* (FP) merupakan suatu informasi data histori yang penting dan dapat digali menggunakan data mining untuk kepentingan kegiatan pemasaran [2]. Isi dari frequent pattern ini adalah pola-pola informasi yang biasanya muncul dalam data transaksi. Metode association rule merupakan salah

satu metode data mining untuk menggali informasi dari frequent pattern yang berguna untuk mengidentifikasi hubungan kesamaan antara item-item.

Pengaplikasian ini dikerjakan di PT. Agro Express Indonesia, merupakan perusahaan yang berkecimpung di bidang penjualan aneka buah-buahan segar yang berdiri sejak tahun 2018. Data yang dipakai pada penelitian ini adalah data transaksi penjualan bulan Oktober hingga November tahun 2018. Hasil analisa mode frekuensi tertinggi adalah kombinasi item yang memenuhi syarat minimum support yang ditentukan. Pembentukan pola aturan asosiasi adalah untuk menemukan aturan asosiasi yang memenuhi kepercayaan minimum dari frekuensi tinggi yang ditemukan. Berikut adalah dataset data transaksi penjualan PT. Agro Express Indonesia bulan Oktober dan November 2018 yang terdiri dari No.Invoice (nomor transaksi) dan nama buah-buahan yang dijual.

NO INVOICE	Almond	Alpokas	Anggur	Apel	Asam Jagruk	Arum Jambu	Blackberry	Buah Naga	Buah Tin	Cherry Australia	Chia Seed	Coconut	Durian Mentan	Grape Grape	Jambu Kuning	Jeruk	Kiwi	Kurma	Lenan	Longkong Bangkok	Madaga	Madras Garden	Nanas Mali	Nanas Australia	Peach Australia	Pisang	Pisang Carondek	Plum	Sakai Fondah
6023	N	N	Y	Y	Y	Y	Y	N	Y	N	Y	N	Y	Y	Y	Y	Y	N	Y	N	Y	N	Y	N	Y	Y	Y	Y	N
6029	N	N	Y	Y	Y	Y	Y	N	Y	N	Y	N	Y	Y	Y	Y	Y	N	Y	N	Y	N	Y	N	Y	Y	Y	Y	N
6030	Y	N	N	Y	Y	Y	Y	N	Y	N	Y	N	Y	Y	Y	Y	Y	N	Y	N	Y	N	Y	N	Y	Y	Y	Y	N
6034	Y	N	N	Y	Y	N	Y	N	Y	Y	Y	Y	N	Y	N	Y	N	N	Y	Y	Y	Y	N	N	N	Y	N	Y	N
6035	Y	N	Y	Y	Y	N	Y	Y	N	Y	Y	Y	N	Y	N	Y	N	Y	N	Y	Y	Y	N	N	N	Y	N	Y	N
6037	Y	N	Y	N	Y	N	Y	Y	N	Y	Y	Y	N	Y	N	Y	N	Y	N	Y	Y	Y	N	N	N	Y	N	Y	N
6039	Y	N	Y	Y	Y	N	Y	Y	N	Y	Y	Y	N	Y	N	Y	N	Y	N	Y	Y	Y	N	N	N	Y	N	Y	N
6042	Y	N	Y	Y	Y	N	Y	N	Y	Y	Y	Y	N	Y	N	Y	N	N	Y	Y	Y	Y	N	N	N	Y	N	Y	N
6043	Y	N	Y	Y	Y	Y	Y	N	N	Y	Y	Y	N	Y	N	Y	N	N	Y	Y	Y	Y	N	N	N	Y	N	Y	N
6044	Y	N	N	Y	Y	Y	Y	N	N	Y	Y	Y	N	Y	N	Y	N	N	N	Y	Y	Y	N	N	N	Y	N	Y	N
6045	Y	N	N	Y	Y	N	Y	Y	N	Y	Y	Y	N	Y	N	Y	N	Y	N	Y	Y	Y	N	N	N	Y	N	Y	N
6046	Y	N	N	Y	Y	N	Y	Y	N	Y	Y	Y	N	Y	N	Y	N	Y	N	Y	Y	Y	N	N	N	Y	N	Y	N
6048	Y	N	N	Y	Y	N	Y	Y	N	Y	Y	Y	Y	N	Y	N	Y	N	Y	Y	Y	Y	N	N	N	Y	N	Y	N
6050	Y	Y	N	N	Y	N	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	N	Y	Y	Y	Y	Y	N
6054	Y	Y	N	N	Y	N	Y	Y	Y	N	N	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	N	N	Y	Y	Y	Y	N
6070	Y	Y	N	N	Y	N	Y	Y	Y	N	N	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	N	Y	Y	Y	Y	Y	N
6079	Y	Y	N	N	Y	N	Y	Y	Y	N	N	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	N	N	Y	Y	Y	Y	N
6083	Y	Y	N	N	Y	N	Y	Y	Y	N	N	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	N	N	Y	Y	Y	Y	N

Gambar 1.1 Sampel Data Transaksi Penjualan PT. Agro Express Indonesia

Pada penelitian sebelumnya, sudah dilakukan penelitian untuk membuat rekomendasi produk dengan algoritma apriori yang dilakukan oleh Ariefana Ria Riszky dan Mujiono Sadikin dengan judul “Data Mining Menggunakan Algoritma Apriori untuk Rekomendasi Produk bagi Pelanggan” [2]. Pada penelitian sebelumnya, terdapat kekurangan pada bagian saran yang dihasilkan, dimana sarannya itu memiliki nilai kekuatan aturan asosiasi yang kecil yaitu 0.352991.

Sehingga tujuan penelitian ini adalah untuk memperbaiki hasil dari penelitian sebelumnya dengan memanfaatkan parameter *tuning* dan juga membandingkan algoritma apriori tersebut dengan *FP-Growth* berdasarkan dari aturan asosiasi yang dihasilkan dan kecepatan pemrosesannya, sehingga mendapatkan algoritma yang paling optimal untuk kasus ini. Untuk tolak ukur perbandingan pada kasus ini yaitu

nilai kekuatan aturan asosiasinya, dimana nilai kekuatan aturan asosiasi itu adalah nilai dari seberapa bagus atau optimalnya aturan valid yang terbentuk oleh suatu algoritma.

Beberapa referensi mengatakan bahwa algoritma FP-Growth itu memiliki nilai kekuatan aturan asosiasi yang lebih besar dibandingkan dengan algoritma Apriori. Dari literatur H. Maulidiya and A. Jananto[3], dengan menggunakan 1805 data, nilai min.support 6%, dan nilai min.confidence 1% menghasilkan nilai kekuatan aturan asosiasi apriori sebesar 0.034326834 dan nilai kekuatan aturan asosiasi fp-growth sebesar 0.198149026. Dari literatur Anthony dan Christofer[4], dengan menggunakan 131 data, nilai min.support 30%, dan nilai min.confidence 60% menghasilkan nilai kekuatan aturan asosiasi apriori sebesar 0.212746294 dan nilai kekuatan aturan asosiasi fp-growth sebesar 0.221762002. Dari literatur Ulya Anisatur Rosyidah dan Hardian Oktavianto[5], dengan menggunakan 670 data, nilai min.support 1%, dan nilai min.confidence 80% menghasilkan nilai kekuatan aturan asosiasi apriori sebesar 0.150149254 dan nilai kekuatan aturan asosiasi fp-growth sebesar 0.203731343. Dari literatur Rizky Mei Anggraeni[6], dengan menggunakan 827 data, nilai min.support 0.4%, dan nilai min.confidence 0.5% menghasilkan nilai kekuatan aturan asosiasi apriori sebesar 0.302443 dan nilai kekuatan aturan asosiasi fp-growth sebesar 0.302443.

Berdasarkan dari beberapa referensi diatas, dalam banyak kasus, algoritma Apriori memiliki kekuatan aturan asosiasi yang lebih kecil dibandingkan dengan algoritma FP-Growth. Sehingga dilakukan perbandingan antara algoritma Apriori yang digunakan dalam penelitian sebelumnya dengan algoritma FP-Growth terhadap kasus ini, dengan asumsi bahwa algoritma FP-Growth kemungkinan besar akan lebih baik dari algoritma Apriori dalam hal kekuatan aturan asosiasi dari aturan yang dihasilkan untuk kasus perekomendasi produk ini.

1.2 Rumusan Masalah

Berdasarkan dari latar belakang di atas, maka dirumuskan permasalahannya yaitu saran yang dihasilkan oleh penelitian sebelumnya masih kurang optimal dan algoritma Apriori ini memiliki kelemahan dari segi kekuatan aturan asosiasi dimana

kekuatan aturan asosiasi dari aturan yang dihasilkan itu lebih kecil atau lebih lemah dibandingkan algoritma FP-Growth.

1.3 Maksud dan Tujuan

Berdasarkan permasalahan di atas, maka maksud dari penelitian ini adalah untuk menentukan algoritma mana yang paling optimal dari segi kekuatan aturan asosiasi dengan menggunakan parameter tuning untuk kasus Rekomendasi Produk bagi Pelanggan ini. Adapun tujuan penelitian ini antara lain :

1. Menentukan nilai kekuatan aturan asosiasi dari aturan yang dihasilkan oleh algoritma Apriori dan FP-Growth dengan menggunakan parameter tuning pada kasus ini.
2. Mendapatkan algoritma yang paling optimal dalam membentuk aturan asosiasi yang memiliki nilai kekuatan aturan asosiasi paling besar yang akan digunakan sebagai acuan untuk membuat rekomendasi produk pada kasus ini.

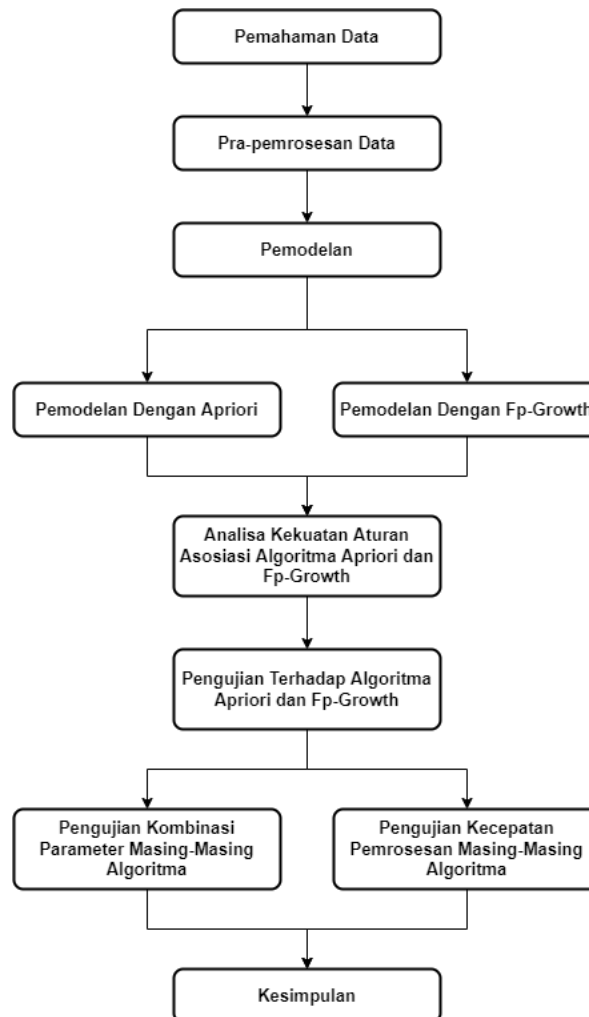
1.4 Batasan Masalah

Permasalahan yang ada cukup luas, sehingga akan dibatasi permasalahannya dalam hal-hal yang diperlukan saja agar pembahasan tidak menyimpang dan lebih terarah. Adapun batasan masalahnya adalah :

1. Analisa dalam penelitian ini berfokus untuk membandingkan kekuatan aturan asosiasi dari aturan yang dihasilkan oleh algoritma Apriori dan FP-Growth dengan menggunakan parameter tuning pada kasus perekomendasi produk ini
2. Data yang digunakan merupakan data PT. Agro Express Indonesia yang menjual beraneka macam buah-buahan segar
3. Sumber data berasal dari data transaksi penjualan bulan Oktober dan November tahun 2018
4. Menggunakan parameter tuning untuk mendapatkan nilai support yang paling optimal
5. Menggunakan parameter tuning untuk mendapatkan nilai confidence yang paling optimal

1.5 Metodologi Penelitian

Metodologi yang digunakan adalah metodologi penelitian terapan dengan tahapan sebagai berikut :



Gambar 1.2 Metodologi Penelitian

1. Pemahaman Data

Pemahaman data adalah tahapan untuk mengumpulkan data dan memahami data yang didapatkan agar bisa masuk kedalam proses data mining. Pengumpulan data yang digunakan dalam penelitian ini yaitu dengan meminta data dari penelitian sebelumnya kepada penulis penelitian sebelumnya yang dimana data tersebut sudah ada di website jurnal dari penelitian tersebut.

2. Prapemrosesan Data

Pada tahap prapemrosesan data ini, mempersiapkan data yang diperoleh pada tahap pengumpulan data agar format dan isinya memenuhi kebutuhan proses data mining. Prapemrosesan data dapat memperbaiki data masukan dan juga meningkatkan akurasi dan efisiensi dari hasil data mining [7]. Data yang diperoleh telah diolah terlebih dahulu dan hanya perlu dilakukan transformasi data sedikit agar sesuai dengan penelitian yang akan dilakukan. Ada 468 data. Setiap data terdiri dari 30 atribut, atribut pertama adalah nomor invoice, dan 29 atribut lainnya adalah nama produk. Berikut adalah data hasil prapemrosesan data.

NO. INVO ICE	Almo nd	Alpu kat	Angg ur	Apel	Asam Bang kok	Avom ango	Blueb erry Arge ntina	Buah Naga Tin	Buah Tin	Cher ry Austr alia	Chia Seed	Deli ma	Durian Mont hong Froze n	Grap eOla	Jamb u Krist al	Jeruk	Kiwi	Kurma	Lemo n	Leng keng Bang kok	Mang ga	Melo n Golde n	Nana s Honi	Necta rin Austr alia	Peac h Austr alia	Pear	Pisan g Cave ndisk	Plum	Salak Pond oh
6423	0	0	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	0	1	0	1	0	1	1	1	1	0
6429	0	0	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	0	1	0	1	0	1	1	1	1	0
6430	1	0	0	1	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	0	1	0	1	0	1	1	1	1	0
6434	1	0	0	1	1	0	1	0	1	1	1	1	0	1	0	1	0	0	1	1	1	1	0	0	0	1	0	1	0
6435	1	0	1	1	1	0	1	1	0	1	1	1	0	1	0	1	0	1	0	1	1	1	0	0	0	1	0	1	0
6437	1	0	1	0	1	0	1	1	0	1	1	1	0	1	0	1	0	1	0	1	1	1	1	0	0	1	0	1	0
6439	1	0	1	0	1	0	1	1	0	1	1	1	0	1	0	1	0	1	0	1	1	1	0	0	0	1	0	1	0
6442	1	0	1	1	1	0	1	0	1	1	1	1	0	1	0	1	0	0	1	1	1	1	0	0	0	1	0	1	0
6443	1	0	1	1	1	1	1	0	0	1	1	1	0	1	0	1	0	0	0	1	1	1	0	0	0	1	0	1	0
6444	1	0	0	1	1	1	1	0	0	1	1	1	0	1	0	1	0	0	0	1	1	1	0	0	0	1	0	1	0

Gambar 1.3 Sampel Data Transaksi yang Dilakukan Transformasi Data

3. Pemodelan dengan Algoritma Apriori

a. Analisis Pola Frekuensi Tinggi

Dalam tahap analisis pola frekuensi tinggi ini, ditentukan kombinasi-kombinasi item yang memenuhi syarat nilai minimum support yang ada dalam seluruh transaksi. Untuk tujuan ini, algoritma apriori digunakan untuk membentuk model data mining untuk menemukan kombinasi itemset dengan nilai frekuensi tertentu. Nilai support ditentukan sesuai dengan batasan masalah yaitu 85%.

b. Pembentukan Association Rule

Setelah frekuensi tinggi dari itemset sudah diperoleh, maka dilanjutkan dengan mencari itemset yang memenuhi syarat minimum confidence dengan menghitung nilai confidence dari seluruh itemset yang sudah memenuhi nilai minimum support. Saat menghitung nilai confidence,

dilakukan penukaran item, seperti kombinasi dua item, yaitu $A \rightarrow B$, lalu balikkan ke $B \rightarrow A$. Kombinasi 3 itemset yang sama yaitu $A, B \rightarrow C$, kemudian dibalik menjadi $A, C \rightarrow B$ dan $B, C \rightarrow A$. Nilai support dari setiap itemset mungkin akan tetap sama, tetapi ada kemungkinan ada itemset-itemset yang akan memiliki nilai confidence yang berbeda. Hal ini untuk mengetahui itemset mana yang memenuhi syarat minimum confidence.

c. Uji Lift Ratio Pada Aturan

Tahap uji lift ratio menghasilkan suatu ukuran untuk menguji kevalidan dari aturan-aturan asosiasi yang terbentuk dan mengetahui kekuatan aturan asosiasi yang terbentuk. Pada tahap ini, untuk mempelajari kevalidan aturan, produk A dibeli bersamaan dengan produk B. Jika nilai lift ratio lebih besar dari 1, maka aturan tersebut dikatakan valid. Artinya aturan tersebut dapat dijadikan sebagai acuan untuk membuat rekomendasi produk.

4. Pemodelan dengan Algoritma FP-Growth

- Menyatukan semua barang dalam setiap transaksi
- Menentukan frekuensi setiap item dari keseluruhan transaksi
- Batasi setiap item dengan support count, dimana item yang memiliki nilai frekuensi lebih kecil dari nilai support count akan dihapus
- Pembangunan Tree berdasarkan per transaksi dengan item yang telah dibatasi
- Tahap Pembangkitan Conditional Pattern Base
- Tahap Pembangkitan Conditional FP-Tree
- Tahap Pencarian Frequent Itemset
- Pengujian data yang telah dihasilkan berupa pola hubungan kombinasi antar item dan association rules sesuai dengan Algoritma FP-Growth
- Menentukan nilai support
- Menghitung nilai confidence

- Mengklasifikasi nilai confidence setiap itemset yang memenuhi syarat confidence minimum
- Melakukan pengajian kevalidan aturan untuk setiap kombinasi itemset. Suatu aturan dapat dikatakan valid jika nilai dari lift ratio aturan tersebut lebih dari 1
- Aturan yang sudah valid akan digunakan sebagai acuan dalam membuat rekomendasi produk

5. Analisa Kekuatan Aturan Asosiasi Algoritma Apriori dan Fp-Growth

Dari hasil analisa kedua algoritma tersebut terhadap kasus Perekomendasi Produk Bagi Pelanggan di PT. Agro Express Indonesia ini, ditampilkan hasil dari setiap algoritma lalu dibandingkan algoritma mana yang memiliki hasil yang paling optimal dari segi kekuatan aturan asosiasinya.

6. Pengujian Kombinasi Parameter Masing-Masing Algoritma

Pada tahap pengujian ini, dilakukan parameter tuning untuk masing-masing algoritma untuk mendapatkan hasil yang paling optimal. Lalu setelah kedua algoritma tersebut mendapatkan hasil nilai kekuatan aturan asosiasi yang paling optimal, akan dibandingkan hasil tersebut untuk menentukan algoritma mana yang paling optimal untuk kasus perekomendasi produk ini.

7. Pengujian Kecepatan Pemrosesan Masing-Masing Algoritma

Pada tahap pengujian kecepatan pemrosesan, dilakukan percobaan terhadap dua device yang berbeda untuk menjalankan kedua algoritma tersebut dan diukur kecepatan pemrosesan dari kedua algoritma tersebut pada dua device. Lalu setelah kedua algoritma tersebut mendapatkan hasil kecepatan pemrosesan di dua device tersebut, akan dibandingkan hasil kecepatan tersebut untuk menentukan algoritma mana yang memiliki kecepatan pemrosesan yang paling cepat untuk kasus perekomendasi produk ini.

8. Kesimpulan

Menuliskan kesimpulan yang didapatkan dari penelitian ini.

1.6 Sistematika Penulisan

Sistematika penulisan disusun untuk memberikan gambaran secara umum mengenai permasalahan dan pemecahannya dari penelitian ini. Sistematika penulisan penelitian ini adalah sebagai berikut :

BAB 1 PENDAHULUAN

Bab ini membahas mengenai latar belakang masalah, rumusan masalah, maksud dan tujuan, batasan masalah, dan metodologi penelitian yang digunakan sebagai acuan dalam tahapan penelitian yang akan dilakukan, serta sistematika penulisan untuk menjelaskan pokok-pokok pembahasannya.

BAB 2 TINJAUAN PUSTAKA

Bab ini menjelaskan mengenai penelitian sebelumnya, objek dari penelitian, konsep-konsep dasar dan juga teori-teori pembangunan sistem yang berkaitan dengan penelitian yang akan dilakukan. Dan juga semua hal yang berguna untuk penelitian ini serta tinjauan terhadap penelitian-penelitian lain yang serupa dan pernah dilakukan sebelumnya.

BAB 3 ANALISIS DAN PERANCANGAN SISTEM

Bab ini berisi analisa masalah dari data hasil penelitian dan proses perancangan sistem yang dibangun sesuai dengan analisis yang dilakukan pada penelitian ini.

BAB 4 IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini berisi mengenai implementasi dan analisis dari perancangan sistem yang telah dilakukan, serta melakukan uji coba terhadap sistem yang telah dibangun sesuai dengan tahapan yang telah dilakukan, dan menampilkan hasil uji coba yang dilakukan terhadap sistem yang dibuat.

BAB 5 KESIMPULAN DAN SARAN

Pada bab ini berisi tentang kesimpulan dari hasil yang telah didapatkan di dalam penelitian tersebut. Juga saran-saran yang berguna untuk mengembangkan sistem tersebut kedepannya.