



**PERBANDINGAN KINERJA ALGORITMA *APRIORI* DAN *FP-GROWTH* TERHADAP *MARKET BASKET ANALYSIS*
PADA DATA PENJUALAN *BAKERY***

SKRIPSI

Oleh:

JULWENTY KARNI GULO

20211000043

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS BUDDHI DHARMA

TANGERANG

2025



**PERBANDINGAN KINERJA ALGORITMA *APRIORI* DAN *FP-GROWTH* TERHADAP *MARKET BASKET ANALYSIS*
PADA DATA PENJUALAN *BAKERY***

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh Gelar Sarjana pada
Program Studi Teknik Informatika Fakultas Sains dan Teknologi**

**Universitas Buddhi Dharma
Jenjang Pendidikan Strata 1**

Oleh:

JULWENTY KARNI GULO

20211000043

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS BUDDHI DHARMA

TANGERANG

2025

LEMBAR PERSEMBAHAN

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa atas segala rahmat, kasih, dan karunia-Nya sehingga saya dapat menyelesaikan skripsi ini dengan baik.

Dengan segala kerendahan hati, saya mempersembahkan karya sederhana ini kepada:

1. Bapak Ebeni Gulo dan Ibu Suriani Gulo tercinta yang telah membesarkan aku dan selalu membimbing, mendukung, memotivasi, memberi apa yang terbaik bagiku serta selalu mendoakan aku untuk meraih kesuksesanku.
2. Adik-adikku Julpekat Wirawan Gulo, Jultry Fictoria Gulo, Inez Yuana Gulo dan Geovan Patriot Gulo yang telah memberikan dukungan semangat serta dorongan yang senantiasa diberikan.
3. Dosen pembimbing Ibu Indah Fenriana, S.Kom.,M.Kom dan seluruh dosen Program Studi Teknik Informatika Universitas Buddhi Dharma yang telah memberikan ilmu, bimbingan, serta arahan selama masa perkuliahan.
4. Teman-teman Seperjuangan Natasya Aprilia, Adisty Maharani, dan yang lain, yang telah berjuang bersama, berbagi cerita, tawa, dan semangat selama masa kuliah hingga terselesaikannya skripsi ini.
5. Andri Fernandes Mabilehi partner yang selalu hadir, mendukung, menemani, dan selalu menyemangati saya, Terima kasih sudah selalu ada.

Semoga skripsi ini dapat memberikan manfaat dan menjadi awal yang baik dalam perjalanan kehidupan saya selanjutnya.

LEMBAR PERNYATAAN KEASLIAN SKRIPSI

Yang bertanda tangan di bawah ini.

NIM	: 20211000043
Nama	: Julwenty Karni Gulo
Jenjang Studi	: Strata I
Program Studi	: Teknik Informatika
Peminatan	: Data Science

Dengan ini saya menyatakan bahwa:

1. Skripsi ini adalah asli dan belum pernah diajukan untuk mendapat gelar akademik Sarjana atau kelengkapan studi, baik di Universitas Buddhi Dharma maupun di Perguruan Tinggi lainnya.
2. Skripsi ini saya buat sendiri tanpa bantuan dari pihak lain, kecuali arahan dosen pembimbing.
3. Dalam Skripsi ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dan dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan daftar pustaka.
4. Dalam Skripsi ini tidak terdapat pemalsuan (kebohongan), seperti buku, artikel, jurnal, data sekunder, pengolahan data, dan pemalsuan tanda tangan dosen atau Ketua Program Studi Universitas Buddhi Dharma yang dibuktikan dengan keasliannya.
5. Lembar pernyataan ini saya buat dengan sesungguhnya, tanpa paksaan dan apabila dikemudian hari atau pada waktu lainnya terdapat penyimpangan dan ketidak benaran dalam pernyataan ini, saya bersedia menerima sanksi akademik berupa pencabutan gelar akademik yang telah saya peroleh karena Skripsi ini serta sanksi lainnya sesuai dengan peraturan dan norma yang berlaku.

Tangerang, 05-08-2025

Yg membuat pernyataan,


Julwenty Karni Gulo

20211000043

LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan di bawah ini.

NIM	: 20211000043
Nama	: Julwenty Karni Gulo
Jenjang Studi	: Strata I
Program Studi	: Teknik Informatika
Peminatan	: Data Science

Dengan ini menyetujui untuk memberikan ijin kepada pihak Universitas Buddhi Dharma, Hak Bebas Royalti Non – Eksklusif (*Non-exclusive Royalty-Free Right*) atas karya ilmiah kami yang berjudul: “Perbandingan Kinerja Algoritma Apriori dan FP-Growth Terhadap *Market Basket Analysis* Pada Data Penjualan *Bakery*”, beserta alat yang diperlukan (apabila ada). Dengan Hak Bebas Royalti Non – Eksklusif ini pihak Universitas Buddhi Dharma berhak menyimpan, mengalih-media atau format-kan, mengelolanya dalam pangkalan data (*database*), mendistribusikannya, dan menampilkan atau mempublikasikannya di *internet* atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis pertama atau pencipta karya ilmiah tersebut. Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Universitas Buddhi Dharma, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Tangerang, 05/08-2025

Yg membuat pernyataan,


Julwenty Karni Gulo

20211000043

Dibuat Oleh :

NIM 20211000043

Nama : Julwenty Karni Gulo

Telah disetujui untuk dipertahankan di hadapan Tim Penguji Ujian

Komprehensif

Program Studi Teknik Informatika

Peminatan Data Science

Tahun Akademik 2024/2025

Disahkan Oleh,

Tangerang, 15-07-2025

Pembimbing,



LEMBAR PENGESAHAN SKRIPSI
PERBANDINGAN KINERJA ALGORITMA APRIORI DAN FP-
GROWTH TERHADAP *MARKET BASKET ANALYSIS*
PADA DATA PENJUALAN *BAKERY*

Dibuat Oleh:

NIM : 20211000043

Nama : Julwenty Karni Gulo

Telah disetujui untuk dipertahankan di hadapan Tim Penguji Ujian
Komprehensif

Program Studi Teknik Informatika

Peminatan Data Science

Tahun Akademik 2024/2025

Disahkan Oleh,

Tangerang, 05-08-2025

Dekan,



Dr. Yakub, M.Kom., M.M.
NUPTK. 1836747648130172

Ketua Program Studi,



Hartana Wijaya, M.Kom.
NUPTK. 3844759660131192

LEMBAR PENGESAHAN TIM PENGUJI

Nama : Julwenty Karni Gulo
NIM : 20211000043
Fakultas : Sains dan Teknologi
Judul Skripsi : Perbandingan Kinerja Algoritma Apriori dan FP-Growth
Terhadap Market Basket Analysis Pada Data Penjualan Bakery

Dinyatakan LULUS setelah mempertahankan di depan Tim Penguji pada hari Selasa, 05-08-2025.

Nama Penguji: Tanda Tangan:

Ketua Sidang : Edy, S.T., M.Kom
NUPTK. 7560760661130203

Penguji I : Hartana Wijaya, M.kom
NUPTK. 3844759660131192

Penguji II : Indah Fenriana, M.Kom
NUPTK. 6538766667230262

Mengetahui,

Dekan Fakultas Sains dan Teknologi

Dr. Yakub, M.Kom., M.M.
NUPTK. 1836747648130172

KATA PENGANTAR

Dengan mengucapkan puji syukur kepada Tuhan Yang Maha Esa, atas rahmat dan karunia-Nya, penulis dapat menyusun dan menyelesaikan Skripsi ini yang berjudul **“PERBANDINGAN KINERJA ALGORITMA APRIORI DAN FP-GROWTH TERHADAP MARKET BASKET ANALYSIS PADA DATA PENJUALAN BAKERY”**. Pembuatan Skripsi ini bertujuan untuk memenuhi salah satu syarat kelengkapan dalam menyelesaikan program pendidikan Strata 1 pada Program Studi Teknik Informatika di Universitas Buddhi Dharma. Dalam proses penyusunan Skripsi ini, penulis memperoleh banyak bantuan dan dukungan, baik secara moril maupun materil, dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Ibu Dr. Limajatini, S.E., M.M., B.K.P. sebagai Rektor Universitas Buddhi Dharma.
2. Bapak Dr. Yakub, M.Kom., M.M., sebagai Dekan Fakultas Sains dan Teknologi.
3. Bapak Hartana Wijaya, M.Kom., sebagai Ketua Program Studi Teknik Informatika.
4. Ibu Indah Fenriana, S.Kom., M.Kom sebagai pembimbing yang telah memberikan bantuan dan dukungan dalam penyelesaian Skripsi ini.
5. Orang tua dan keluarga yang selalu memberikan dukungan, baik moril maupun materiil.
6. Teman-teman yang selalu memberikan semangat dan bantuan.

Tak lupa juga, penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan kontribusi, meskipun tidak dapat disebutkan satu per satu, yang turut mewujudkan terselesainya penulisan ini. Penulis menyadari bahwa Skripsi ini masih jauh dari sempurna, oleh karena itu, kritik dan saran yang bersifat membangun sangat penulis harapkan demi perbaikan di masa yang akan datang.

Akhir kata, semoga Skripsi ini dapat memberikan manfaat bagi penulis khususnya, dan bagi para pembaca yang berminat pada umumnya.

Tangerang, 05 Agustus 2025

Penulis

ABSTRAK

Perkembangan teknologi informasi telah mendorong berbagai sektor bisnis untuk memanfaatkan data sebagai aset berharga, salah satunya melalui analisis pola pembelian konsumen dengan metode Market Basket Analysis (MBA). Penelitian ini bertujuan untuk membandingkan kinerja algoritma Apriori dan FP-Growth dalam menganalisis data transaksi penjualan pada toko bakery menggunakan dataset Bakery.csv. Metode penelitian yang digunakan meliputi tahap pra-pemrosesan data, penerapan algoritma Apriori dan FP-Growth, pengujian berdasarkan kombinasi parameter, serta evaluasi kinerja berdasarkan waktu eksekusi, penggunaan memori, dan jumlah aturan asosiasi yang dihasilkan. Hasil penelitian menunjukkan bahwa pada kombinasi support 0,01 & confidence 0,5, kedua algoritma menghasilkan 11 aturan asosiasi, dengan Apriori lebih cepat (0,0543 detik, 6,02 MB) dibanding FP-Growth (4,0944 detik, 10,38 MB). Pada kombinasi support 0,02 & confidence 0,6, keduanya menghasilkan 1 aturan, dengan Apriori (0,0191 detik, 2,50 MB) tetap lebih efisien dibanding FP-Growth (2,1680 detik, 10,38 MB). Sedangkan pada kombinasi support 0,05 & confidence 0,7, kedua algoritma tidak menghasilkan aturan, namun Apriori tetap unggul dari segi waktu (0,0204 detik vs 0,3369 detik) dan penggunaan memori (0,57 MB vs 10,38 MB). Berdasarkan hasil tersebut, dapat disimpulkan bahwa algoritma Apriori lebih efisien dalam hal waktu eksekusi dan penggunaan memori pada dataset skala menengah seperti Bakery.csv. Hasil analisis asosiasi dapat dimanfaatkan untuk strategi pemasaran berbasis data, seperti cross-selling dan rekomendasi produk. Penelitian ini merekomendasikan penggunaan algoritma Apriori untuk dataset serupa, serta pengembangan antarmuka pengguna yang lebih interaktif dan pengoptimalan algoritma untuk dataset berskala besar.

Kata kunci : *Algoritma Apriori, Dataset Bakery, Data mining, FP-Growth, Market Basket Analysis (MBA).*

ABSTRACT

The development of information technology has driven various business sectors to utilize data as a valuable asset, one of which is through analyzing consumer purchasing patterns using the Market Basket Analysis (MBA) method. This study aims to compare the performance of the Apriori and FP-Growth algorithms in analyzing bakery sales transaction data using the Bakery.csv dataset. The research methodology includes data preprocessing, implementation of the Apriori and FP-Growth algorithms, parameter combination testing, and performance evaluation based on execution time, memory usage, and the number of generated association rules. The results show that with a support value of 0.01 and a confidence value of 0.5, both algorithms produced 11 association rules, with Apriori performing faster (0.0543 seconds, 6.02 MB) compared to FP-Growth (4.0944 seconds, 10.38 MB). With a support value of 0.02 and a confidence value of 0.6, both produced 1 rule, where Apriori (0.0191 seconds, 2.50 MB) remained more efficient than FP-Growth (2.1680 seconds, 10.38 MB). At a support value of 0.05 and a confidence value of 0.7, no rules were generated by either algorithm; however, Apriori still outperformed FP-Growth in terms of execution time (0.0204 seconds vs. 0.3369 seconds) and memory usage (0.57 MB vs. 10.38 MB). Based on these results, it can be concluded that the Apriori algorithm is more efficient in terms of execution time and memory usage for medium-scale datasets such as Bakery.csv. The association analysis results can be utilized for data-driven marketing strategies, such as cross-selling and product recommendations. This study recommends the use of the Apriori algorithm for similar datasets, as well as the development of more interactive user interfaces and algorithm optimization for large-scale datasets.

Keywords: *Algoritma Apriori, Dataset Bakery, Data mining, FP-Growth, Market Basket Analysis (MBA).*

DAFTAR ISI

LEMBAR PERSEMBAHAN	i
LEMBAR PERNYATAAN KEASLIAN SKRIPSI.....	ii
LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH.....	iii
LEMBAR PENGESAHAN PEMBIMBING SKRIPSI.....	iv
LEMBAR PENGESAHAN SKRIPSI.....	v
LEMBAR PENGESAHAN TIM PENGUJI SKRIPSI	vi
KATA PENGANTAR	vii
ABSTRAK.....	viii
<i>ABSTRACT</i>	ix
DAFTAR ISI	x
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR.....	xv
DAFTAR LAMPIRAN.....	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang Penelitian.....	1
1.2 Identifikasi Masalah.....	2
1.3 Ruang Lingkup Penelitian	3
1.4 Tujuan dan Manfaat Penelitian	3
1.4.1 Tujuan Penelitian.....	3
1.4.2 Manfaat Penelitian.....	4
1.5 Sistematika Penulisan	4
BAB II LANDASAN TEORI.....	6
2.1 Teori – Teori yang Digunakan.....	6
2.1.1 Data	6
2.1.2 Informasi	7

2.1.3 Data Mining	9
2.1.4 Knowledge Discovery in Database (KDD)	10
2.1.5 Market Basket Analysis (MBA)	11
2.1.6 Association Rule Mining	12
2.1.7 Algoritma Apriori	14
2.1.8 Algoritma FP-Growth	15
2.1.9 Prediksi/Forecasting	16
2.1.10 Penjualan	17
2.2 Penelitian yang Relevan	18
2.2.1 Comparison of Market Basket Analysis to Determine Consumer Purchasing Patterns Using FP-Growth and Apriori Algorithm	18
2.2.2 Market Basket Analysis Using Apriori and FP Growth Algorithm	19
2.2.3 Evaluating The Performance Of Association Rules In Apriori And FP-Growth Algorithms: Market Basket Analysis To Discoverrules Of Item Combinations	20
2.2.4 Perbandingan Algoritma Apriori dan FP-Growth Terhadap Market Basket Analysis Pada Data Penjualan Toko Roti	20
2.2.5 Rancang Bangun Aplikasi Web untuk Meningkatkan Penjualan melalui Analisis Pola Pembelian	21
2.2.6 Implementasi Algoritma Apriori pada Penjualan Alat Teknik Pertanian	21
2.2.7 Penerapan Algoritma FP-Growth Untuk Menentukan Pola Pembelian Konsumen Pada AHASS Cibadak	22
2.2.8 Implementasi Algoritma FP-Growth Untuk Menemukan Pola Keterkaitan Antara Mata Kuliah Pemrograman Dan Mata kuliah Matematika	23
2.2.9 Implementasi Market Basket Analysis Pada Data Penjualan CV.XYZ Menggunakan Algoritma FP-Growth	23
2.2.10 Market Basket Analysis Of Administrative Patterns Data Of Consumer Purchases Using Data Mining Technology	24
2.2.11 Application of Market Basket Analysis on Beauty Clinic to Increasing Customer's Buying Decision	24

2.2.12 Implementasi <i>Data Mining</i> Menggunakan Algoritma <i>Apriori</i> Dalam Menentukan Persediaan Barang (Studi Kasus: Toko Sinar Harahap).....	25
2.2.13 Implementasi Metode <i>Market Basket Analysis</i> Pada Sistem Penjualan Dan Pemasaran Butik Nazwa Fashion.....	25
2.2.14 <i>Market Basket Analysis</i> Apotek Berbasis <i>Web</i> Menggunakan Metode Algoritma <i>Apriori</i>	26
2.2.15 <i>Comparison of Apriori, Apriori-TID, and FP-Growth Algorithms in Market Basket Analysis at Grocery Stores</i>	26
BAB III METODOLOGI PENELITIAN.....	28
3.1 Analisa Kebutuhan.....	28
3.1.1 Kebutuhan Data.....	28
3.1.2 Kebutuhan <i>Tools</i> dan <i>Software</i>	32
3.1.3 Kebutuhan <i>Library</i>	33
3.2 Pembahasan Metode yang Digunakan.....	34
3.2.1 Tahapan Metode Penelitian.....	34
3.2.2 Penentuan Parameter Minimum <i>Support</i> dan Minimum <i>Confidence</i>	36
3.2.4 <i>Output</i> yang Diharapkan.....	39
3.3 Pembahasan Algoritma yang Digunakan.....	40
3.3.1 Algoritma <i>Apriori</i>	40
3.3.2 Algoritma <i>FP-Growth</i>	45
3.3.3 Analisis Perbandingan Kinerja Algoritma.....	53
3.4 Kerangka Pemikiran.....	55
3.5 Perancangan <i>UML</i>	56
3.6.1 <i>Use Case Diagram</i>	57
3.6.2 <i>Activity Diagram</i>	57
3.6 Perancangan Layar dan Menu.....	59
3.7 Jadwal Penelitian.....	62

BAB IV HASIL DAN PEMBAHASAN	63
4.1 Spesifikasi <i>Hardware</i> dan <i>Software</i>	63
4.1.1 <i>Hardware</i>	63
4.1.2 <i>Software</i>	63
4.2 Tampilan Aplikasi	64
4.2.1 Tampilan Awal atau Menu Beranda Aplikasi	64
4.2.2 Tampilan Menu <i>Dataset</i>	64
4.2.3 Tampilan Menu Parameter	67
4.2.4 Tampilan Menu Analisis.....	68
4.2.5 Tampilan Menu Riwayat Analisis	71
4.3 Penerapan Metode dan Algoritma.....	72
4.4 <i>Black box Testing</i>	74
4.4.1 <i>Black Box Testing</i> Menu Beranda	74
4.4.2 <i>Black Box Testing</i> Menu <i>Dataset</i>	75
4.4.3 <i>Black Box Testing</i> Menu Parameter.....	75
4.4.4 <i>Black Box Testing</i> Menu Analisis	76
4.4.5 <i>Black Box Testing</i> Menu Riwayat Analisis.....	76
4.5 <i>User Acceptance Testing (UAT)</i>	77
BAB V SIMPULAN DAN SARAN	83
5.1 SIMPULAN.....	83
5.2 SARAN.....	85
DAFTAR PUSTAKA	86

DAFTAR TABEL

Tabel 3.1 Keterangan Atribut <i>Dataset Bakery</i>	29
Tabel 3.2 Kebutuhan <i>Tools</i> dan <i>Software</i>	32
Tabel 3.3 <i>Library</i> yang Digunakan.....	33
Tabel 3.4 Data sampel <i>apriori</i>	42
Tabel 3.5 Menghitung <i>Support Apriori</i>	43
Tabel 3.6 <i>Itemset</i> yg memenuhi <i>support</i>	43
Tabel 3.7 Pembentukan Aturan Asosiasi	43
Tabel 3.8 Nilai <i>item</i> Prioritas <i>FP-Growth</i>	46
Tabel 3.9 Mengurutkan <i>item</i> prioritas <i>FP-Growth</i>	46
Tabel 3.10 <i>Frequent Pattern Generate</i>	52
Tabel 3.11 Aturan Asosiasi <i>FP-Growth</i>	52
Tabel 3.12 Hasil Perbandingan <i>Apriori</i> dan <i>FP-Growth</i>	53
Tabel 3.13 Jadwal Penelitian	62
Tabel 4.1 Spesifikasi Hardware	63
Tabel 4.2 Spesifikasi <i>Software</i>	63
Tabel 4.3 Perbandingan Kinerja Algoritma <i>Apriori</i> dan <i>FP-Growth</i>	73
Tabel 4.4 <i>Black Box</i> Menu Beranda	75
Tabel 4.5 <i>Black Box</i> Menu <i>Dataset</i>	75
Tabel 4.6 <i>Black Box</i> Menu Parameter	76
Tabel 4.7 <i>Black Box</i> Menu Analisis.....	76
Tabel 4.8 <i>Black Box</i> Menu Riwayat Analisis	77

DAFTAR GAMBAR

Gambar 3.1 Gambar <i>Dataset</i> Asli.....	28
Gambar 3.2 Top 10 <i>Item</i> Terlaris	31
Gambar 3.3 Jumlah Transaksi per Jam	31
Gambar 3.4 Jumlah Transaksi per Hari	32
Gambar 3.5 Tahapan Metode Penelitian.....	35
Gambar 3.6 Hasil Analisis <i>Apriori</i>	44
Gambar 3.7 Langkah Pertama Pembentukan <i>FP-Tree</i>	47
Gambar 3.8 Langkah Kedua Pembentukan <i>FP-Tree</i>	47
Gambar 3.9 Langkah Ketiga Pembentukan <i>FP-Tree</i>	48
Gambar 3.10 Langkah Keempat Pembentukan <i>FP-Tree</i>	48
Gambar 3.11 Langkah Kelima Pembentukan <i>FP-Tree</i>	49
Gambar 3.12 Langkah Keenam Pembentukan <i>FP-Tree</i>	50
Gambar 3.13 Langkah Ketujuh Pembentukan <i>FP-Tree</i>	50
Gambar 3.14 Langkah Kedelapan Pembentukan <i>FP-Tree</i>	51
Gambar 3.15 Langkah Kesembilan Pembentukan <i>FP-Tree</i>	51
Gambar 3.16 Hasil Analisis <i>FP-Growth</i>	53
Gambar 3.17 Kerangka Pemikiran	56
Gambar 3.18 Use Case Diagram	57
Gambar 3.19 Activity Diagram	58
Gambar 3.20 Rancangan Halaman Beranda.....	59
Gambar 3.21 Rancangan Menu Dataset	60
Gambar 3.22 Rancangan Menu Parameter.....	60
Gambar 3.23 Rancangan Menu Analisis	61
Gambar 3.24 Rancangan Menu <i>Histori</i>	62
Gambar 4.1 Tampilan Menu Beranda Aplikasi	64
Gambar 4.2 Tampilan Menu <i>Upload Dataset</i>	65
Gambar 4.3 Tampilan Setelah <i>Upload Dataset</i>	66
Gambar 4.4 Tampilan Menu Pengaturan Parameter Analisis	67
Gambar 4.5 Tampilan Hasil Analisis	68
Gambar 4.6 Tampilan Hasil <i>Apriori</i>	69
Gambar 4.7 Tampilan Hasil <i>FP-Growth</i>	70
Gambar 4.8 Tampilan Menu Riwayat Analisis.....	71

Gambar 4.9 Tampilan Riwayat Parameter.....	71
Gambar 4.10 Tampilan <i>Tab</i> Detail.....	72



DAFTAR LAMPIRAN

Lampiran 1 Kartu Bimbingan Skripsi	91
Lampiran 2 Hasil Jawaban <i>Google Form</i>	93



BAB I

PENDAHULUAN

1.1 Latar Belakang Penelitian

Perkembangan teknologi informasi yang pesat di era digital saat ini telah mendorong berbagai sektor bisnis untuk memanfaatkan data sebagai aset yang sangat berharga. Salah satu teknik analisis data yang banyak diterapkan dalam dunia bisnis adalah *Market Basket Analysis* (MBA), yang bertujuan untuk menemukan pola hubungan antara produk yang sering dibeli bersamaan oleh konsumen berdasarkan data transaksi penjualan. Dengan menggunakan MBA, perusahaan dapat memperoleh wawasan yang lebih dalam mengenai perilaku konsumen, mengoptimalkan penataan produk di rak toko, serta merancang strategi pemasaran yang lebih efektif dan tepat sasaran (Idris et al., 2022).

Dalam penerapan *Market Basket Analysis*, algoritma *association rule mining* menjadi metode utama untuk menemukan pola asosiasi antar produk. Dua algoritma yang sering digunakan adalah *Apriori* dan *FP-Growth*. Algoritma *Apriori* membangkitkan kandidat secara *iteratif*, sedangkan *FP-Growth* menggunakan struktur pohon (*frequent pattern tree*) untuk menemukan pola *frequent itemset* tanpa perlu membangkitkan kandidat secara eksplisit (Aktavera et al., 2024). Meskipun kedua algoritma ini telah banyak diterapkan, keduanya memiliki kelebihan dan kekurangan yang signifikan. *Apriori* cenderung memerlukan waktu komputasi yang lebih lama pada dataset yang besar, karena proses iterasi yang berulang, sedangkan *FP-Growth* lebih efisien dalam hal waktu komputasi, tetapi membutuhkan lebih banyak memori untuk membangun struktur pohon (Musdalifah & Jananto, 2022).

Penelitian ini menggunakan *dataset* transaksi penjualan *bakery* yang diambil dari *platform Kaggle*. Meskipun berasal dari industri *bakery*, struktur dan karakteristik *dataset* ini serupa dengan data transaksi yang ada di industri ritel dan *e-commerce* secara umum. Oleh karena itu, hasil eksperimen dan perbandingan kinerja yang dilakukan pada *dataset* ini diharapkan dapat memberikan gambaran yang relevan dan aplikatif untuk berbagai sektor bisnis yang juga memanfaatkan MBA, seperti *supermarket*, toko grosir, dan *platform e-commerce*.

Pemilihan algoritma yang tepat untuk melakukan MBA sangat penting, terutama ketika dihadapkan dengan data transaksi yang besar dan kompleks. Penelitian ini bertujuan untuk membandingkan kinerja dua algoritma, *Apriori* dan *FP-Growth*, dalam melakukan MBA pada data transaksi penjualan *bakery*. Kinerja kedua algoritma akan dibandingkan berdasarkan beberapa metrik, seperti waktu komputasi, penggunaan memori, serta kualitas aturan asosiasi yang dihasilkan, dengan menggunakan metrik evaluasi seperti *support*, *confidence*, dan *lift*. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi yang signifikan bagi akademisi, praktisi data science, serta pelaku bisnis dalam memilih algoritma yang paling sesuai untuk analisis pola pembelian konsumen dan menjadi acuan dalam pengembangan strategi pemasaran berbasis data.

1.2 Identifikasi Masalah

Berdasarkan latar belakang sebelumnya, maka permasalahan yang diidentifikasi dalam penelitian ini adalah:

1. Belum diketahui perbedaan kinerja algoritma *Apriori* dan *FP-Growth* dalam menganalisis pola pembelian konsumen pada data transaksi penjualan *bakery*.

2. Belum terdapat informasi yang jelas mengenai algoritma mana yang lebih efisien jika dilihat dari aspek waktu eksekusi, penggunaan memori, dan kualitas aturan asosiasi yang dihasilkan berdasarkan nilai *support*, *confidence*, dan *lift*.
3. Hasil analisis pola pembelian konsumen dari *Market Basket Analysis* pada data penjualan *bakery* belum dimanfaatkan secara optimal untuk mendukung strategi pemasaran, seperti *cross-selling* dan rekomendasi produk.

1.3 Ruang Lingkup Penelitian

Penelitian ini difokuskan pada:

1. Penggunaan dataset *Bakery.csv* dari *Kaggle* yang berisi data penjualan produk *bakery* seperti *bread*, *cake*, dan *pastry*.
2. Penerapan algoritma *Apriori* dan *FP-Growth* untuk analisis pola pembelian konsumen.
3. Evaluasi performa berdasarkan parameter: waktu komputasi, penggunaan memori, dan kualitas aturan yang dihasilkan.
4. Penggunaan nilai minimum *support* dan *confidence* yang telah ditentukan.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Membandingkan efisiensi kedua algoritma berdasarkan waktu eksekusi, penggunaan memori, dan kualitas aturan asosiasi (*support*, *confidence*, dan *lift*).
2. Memberikan rekomendasi algoritma yang optimal untuk penerapan *Market Basket Analysis* pada industri *bakery*.

3. Menyajikan gambaran pemanfaatan hasil *Market Basket Analysis* untuk mendukung strategi pemasaran berbasis data.

1.4.2 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah:

1. Memberikan referensi ilmiah bagi akademisi dan praktisi dalam memilih algoritma data mining yang efisien untuk analisis pola pembelian konsumen.
2. Menjadi acuan bagi pelaku bisnis atau pengembang sistem informasi penjualan dalam memilih algoritma yang tepat untuk *Market Basket Analysis*.
3. Memberikan wawasan kepada pelaku usaha mengenai pemanfaatan hasil *Market Basket Analysis* dalam perancangan strategi pemasaran yang lebih efektif.

1.5 Sistematika Penulisan

Adapun sistematika penulisan laporan ini adalah sebagai berikut:

BAB I PENDAHULUAN

Berisi latar belakang penelitian, identifikasi masalah, ruang lingkup penelitian, tujuan dan manfaat penelitian, serta sistematika penulisan.

BAB II LANDASAN TEORI

Membahas teori-teori yang mendukung penelitian, termasuk konsep *Market Basket Analysis*, algoritma *Apriori*, algoritma *FP-Growth*, serta penelitian terdahulu yang relevan.

BAB III METODOLOGI PENELITIAN

Menjelaskan tahapan penelitian, analisis kebutuhan, metode dan algoritma yang digunakan, kerangka berpikir, *flowchart* sistem, serta rancangan antarmuka dan *database*.

BAB IV HASIL DAN PEMBAHASAN

Menyajikan spesifikasi *hardware* dan *software*, tampilan program, penerapan metode dan algoritma, hasil pengujian (*blackbox* dan UAT), serta analisis hasil penelitian.

BAB V SIMPULAN DAN SARAN

Menyimpulkan hasil penelitian dan memberikan saran untuk penelitian lebih lanjut.



BAB II

LANDASAN TEORI

2.1 Teori – Teori yang Digunakan

2.1.1 Data

Data merupakan elemen fundamental dalam setiap proses penelitian, terutama dalam pendekatan kuantitatif yang menitikberatkan pada pengumpulan dan analisis data *numerik* untuk menghasilkan kesimpulan yang bersifat objektif dan dapat diuji secara statistik. (Arhami, M. & Nasir, 2020), data dapat dipahami sebagai bahan mentah berupa fakta, angka, atau simbol yang belum memiliki makna sampai melalui proses pengolahan dan interpretasi menjadi informasi. Oleh karena itu, data berperan sebagai dasar utama dalam membangun pengetahuan baru yang relevan dengan fokus penelitian.

Secara umum, data dikategorikan ke dalam dua bentuk utama, yaitu data kualitatif dan data kuantitatif. Data kualitatif mengacu pada informasi yang diekspresikan dalam bentuk narasi, teks, gambar, atau rekaman suara yang tidak dapat diukur secara *numerik*, dan biasanya digunakan untuk menggambarkan atau memahami fenomena secara mendalam melalui pendekatan interpretatif. Sebaliknya, data kuantitatif berwujud angka atau nilai *numerik* yang dapat dihitung dan dianalisis secara statistik, sehingga cocok digunakan untuk menguji hipotesis atau menemukan pola dalam himpunan data yang besar.

Berdasarkan asal usulnya, data juga dapat diklasifikasikan menjadi data primer dan data sekunder. Data primer diperoleh secara langsung oleh peneliti melalui teknik pengumpulan data seperti observasi, wawancara, survei, atau

eksperimen. Sementara itu, data sekunder merupakan data yang telah dikumpulkan oleh pihak lain dan tersedia dalam bentuk laporan penelitian, artikel jurnal ilmiah, atau dokumen resmi perusahaan.

Dalam konteks penelitian ini, data yang digunakan adalah data transaksi penjualan ritel pada sebuah toko roti (*bakery*). Data tersebut mencakup informasi pembelian konsumen seperti daftar produk yang dibeli, jumlah item, waktu transaksi, serta atribut relevan lainnya. Karena berbentuk numerik, data ini termasuk dalam kategori kuantitatif dan dapat digunakan untuk melakukan analisis keranjang pasar (*market basket analysis*) guna mengidentifikasi pola pembelian pelanggan. Hasil dari analisis ini dapat digunakan untuk mendukung pengambilan keputusan strategis, seperti pengaturan tata letak produk, kebijakan promosi, dan perencanaan strategi pemasaran (Prawira et al., 2020).

2.1.2 Informasi

Informasi merupakan hasil dari serangkaian proses pengolahan data yang menjadikannya bermakna, bernilai, dan bermanfaat bagi penggunaannya. Tidak seperti data mentah yang bersifat deskriptif dan belum terstruktur, informasi tercipta ketika data telah melalui tahap pengorganisasian, pemrosesan, dan interpretasi sehingga dapat memberikan kontribusi nyata dalam proses pengambilan Keputusan. (Anggraeni, I., & Irviani, 2017), mendefinisikan informasi sebagai data yang telah diproses sedemikian rupa sehingga memiliki nilai guna yang lebih tinggi serta mampu mengurangi ketidakpastian dalam proses penilaian atau pengambilan keputusan atas suatu situasi tertentu.

Karakteristik informasi yang berkualitas mencakup sejumlah dimensi penting, antara lain: akurasi, relevansi, ketepatan waktu, kelengkapan,

keterpahaman, verifiabilitas, dan aksesibilitas. Informasi yang akurat mencerminkan kondisi aktual dan terbebas dari kekeliruan. Sementara itu, relevansi merujuk pada tingkat kesesuaian informasi terhadap kebutuhan pengguna dan konteks permasalahan yang dihadapi. Ketepatan waktu sangat krusial karena informasi yang terlambat dapat kehilangan nilai guna strategisnya. Kriteria lainnya, seperti kemampuan untuk diverifikasi dan diakses dengan mudah, juga menjadi indikator kualitas suatu informasi dalam sistem manajemen berbasis data.

Menurut Romney dan Steinbart (2016), nilai sebuah informasi ditentukan oleh selisih antara manfaat yang diperoleh dari penggunaan informasi tersebut dan biaya yang dikeluarkan untuk menghasilkan atau mengaksesnya. Manfaat informasi dapat berupa pengurangan ketidakpastian, peningkatan efektivitas pengambilan keputusan, serta optimalisasi perencanaan dan pengorganisasian aktivitas bisnis.

Dalam konteks penelitian ini, informasi diperoleh melalui analisis terhadap data transaksi penjualan ritel, khususnya dari aktivitas pembelian konsumen di toko *bakery*. Proses analitis ini menghasilkan wawasan berupa pola pembelian yang berulang dan hubungan antarproduk, yang selanjutnya dapat digunakan untuk merumuskan strategi bisnis yang lebih responsif, seperti penempatan produk di etalase, promosi tersegmentasi, serta pengembangan kebijakan pemasaran yang berbasis data.

Dengan demikian, kualitas informasi yang dihasilkan dari pengolahan data transaksi menjadi sangat penting. Informasi tersebut harus memenuhi prinsip-

prinsip dasar kualitas informasi agar dapat mendukung proses pengambilan keputusan secara efektif dan efisien dalam konteks manajemen ritel.

2.1.3 Data Mining

Data mining merujuk pada proses sistematis untuk mengidentifikasi pola tersembunyi, hubungan antarvariabel, dan pengetahuan baru dari himpunan data berukuran besar menggunakan pendekatan statistik, algoritmik, dan komputasional. Tujuan utama dari proses ini adalah untuk mengekstraksi informasi bernilai tambah yang sebelumnya tidak tampak secara eksplisit, guna mendukung pengambilan keputusan yang berbasis data. Dalam terminologi yang lebih luas, *data mining* merupakan bagian integral dari proses *Knowledge Discovery in Databases (KDD)*, yang terdiri atas beberapa tahap berurutan: pembersihan data (*data cleaning*), integrasi data (*data integration*), seleksi data (*data selection*), transformasi data, eksplorasi data (*data mining*), evaluasi pola (*pattern evaluation*), dan visualisasi atau penyajian pengetahuan kepada pengguna akhir (Han Pei et al., 2021).

Menurut (Krisna et al., 2022), *data mining* berfungsi untuk mengidentifikasi informasi tersembunyi yang dapat meningkatkan kualitas keputusan strategis. Output dari proses ini bisa berupa pola asosiatif, prediktif, atau deskriptif, seperti aturan asosiasi (*association rules*), klasifikasi (*classification*), regresi (*regression*), serta segmentasi data melalui klusterisasi (*clustering*). Teknik dan algoritma yang digunakan dipilih berdasarkan karakteristik data serta tujuan analisis yang ingin dicapai.

Dalam penelitian ini, *data mining* diterapkan untuk menganalisis perilaku konsumen berdasarkan data transaksi penjualan ritel dari sebuah toko roti. Dengan

menggunakan algoritma populer seperti *Apriori* dan *FP-Growth*, peneliti mengidentifikasi keterkaitan antar produk yang sering dibeli secara bersamaan—sebuah pendekatan yang dikenal sebagai *market basket analysis*. Pola pembelian ini dapat digunakan untuk menyusun strategi bisnis yang lebih adaptif, seperti optimalisasi tata letak produk, rekomendasi promosi terpadu, dan penyusunan paket penjualan yang sesuai dengan preferensi konsumen.

2.1.4 Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) merupakan proses menyeluruh dalam menemukan pengetahuan baru yang bermanfaat dari kumpulan data yang besar. KDD tidak hanya mencakup proses data mining, tetapi juga serangkaian tahapan lain yang saling berkesinambungan, mulai dari seleksi data, pembersihan data, transformasi data, proses penambangan data (*data mining*), hingga evaluasi dan interpretasi hasil yang diperoleh.

Menurut (Putri. P et al., 2021), KDD membantu dalam menemukan pola-pola tersembunyi yang tidak dapat ditemukan dengan teknik analisis data biasa. Proses KDD memungkinkan organisasi atau peneliti untuk memperoleh wawasan baru yang dapat digunakan dalam pengambilan keputusan strategis, terutama ketika berhadapan dengan data dalam jumlah besar dan kompleks. Secara umum, tahapan utama dalam proses KDD adalah sebagai berikut:

1. Seleksi Data (*Data Selection*): Memilih data yang relevan dari basis data yang tersedia sesuai dengan tujuan analisis.
2. Pembersihan Data (*Data Cleaning*): Menghilangkan data yang tidak konsisten, duplikat, atau tidak lengkap agar data yang digunakan berkualitas baik.

3. Transformasi Data (*Data Transformation*): Mengubah atau mengintegrasikan data ke dalam format yang sesuai untuk proses analisis, misalnya normalisasi atau agregasi data.
4. *Data Mining*: Proses inti KDD, yaitu penerapan metode atau algoritma tertentu untuk menemukan pola, tren, atau hubungan dalam data.
5. Evaluasi dan Interpretasi (*Evaluation and Interpretation*): Menilai pola-pola yang ditemukan untuk memastikan bahwa pola tersebut valid, berguna, dan dapat diinterpretasikan sesuai kebutuhan bisnis atau penelitian.

Dalam penelitian ini, KDD digunakan untuk menemukan pola pembelian konsumen dari data transaksi penjualan ritel. Proses KDD memungkinkan peneliti untuk memperoleh pengetahuan baru yang dapat digunakan untuk meningkatkan strategi pemasaran, penataan produk, dan pengambilan keputusan bisnis lainnya.

2.1.5 Market Basket Analysis (MBA)

Market Basket Analysis (MBA) adalah teknik analisis data yang digunakan untuk menemukan hubungan atau asosiasi antar produk yang sering dibeli secara bersamaan oleh konsumen dalam satu transaksi. Teknik ini sangat populer dalam industri ritel karena dapat membantu bisnis memahami pola belanja konsumen dan merancang strategi pemasaran yang lebih efektif.

MBA bekerja dengan menganalisis data transaksi penjualan untuk mengidentifikasi itemset, yaitu kelompok produk yang sering muncul bersama dalam keranjang belanja pelanggan. Dengan memanfaatkan algoritma seperti *Apriori* atau *FP-Growth*, MBA dapat menghasilkan aturan asosiasi yang menunjukkan kemungkinan keterkaitan antar produk, misalnya: "Jika konsumen membeli roti, maka kemungkinan besar juga membeli mentega".

Menurut (Khedkar & Kumari, 2021), *Market Basket Analysis* membantu bisnis dalam merancang strategi pemasaran yang efektif melalui pemahaman pola belanja konsumen, seperti penempatan produk yang saling terkait secara berdampingan, penawaran promosi *bundling*, atau rekomendasi produk secara otomatis. Informasi yang dihasilkan dari MBA juga dapat digunakan untuk mengoptimalkan pengelolaan inventori, meningkatkan efektivitas promosi silang (*cross-selling*), serta meningkatkan kepuasan dan loyalitas pelanggan.

Selain itu, MBA juga digunakan dalam pengembangan sistem rekomendasi produk, baik di toko fisik maupun *e-commerce*, untuk meningkatkan penjualan dan pengalaman belanja pelanggan. Contoh penerapan MBA adalah fitur "*Frequently Bought Together*" pada situs belanja daring, di mana sistem merekomendasikan produk tambahan yang relevan berdasarkan pola pembelian konsumen sebelumnya.

Dengan demikian, *Market Basket Analysis* menjadi salah satu alat analisis utama dalam *data mining* untuk mendukung pengambilan keputusan bisnis berbasis data di sektor ritel.

2.1.6 Association Rule Mining

Association Rule Mining adalah teknik dalam *data mining* yang bertujuan untuk menemukan aturan-aturan asosiatif (*association rules*) di antara item-item dalam kumpulan data yang besar. Aturan asosiatif ini berbentuk pernyataan "jika-maka" (*if-then*) yang menghubungkan dua atau lebih itemset, misalnya: "Jika konsumen membeli roti, maka kemungkinan besar juga membeli mentega" (Musdalifah & Jananto, 2022).

Setiap aturan asosiatif memiliki dua bagian utama, yaitu *antecedent* (bagian “jika”) dan *consequent* (bagian “maka”). Proses ini sangat penting dalam analisis data transaksi ritel, seperti *market basket analysis*, untuk mengidentifikasi pola pembelian konsumen yang dapat dimanfaatkan dalam strategi bisnis.

Evaluasi kualitas dan kekuatan aturan asosiatif dilakukan menggunakan tiga metrik utama, yaitu *support*, *confidence*, dan *lift*:

1. *Support* mengukur seberapa sering itemset tertentu muncul dalam seluruh data transaksi. *Support* dihitung sebagai rasio jumlah transaksi yang mengandung itemset tersebut terhadap total seluruh transaksi. Nilai *support* yang tinggi menunjukkan pola tersebut sering terjadi dan signifikan dalam dataset.
2. *Confidence* mengukur seberapa sering aturan asosiatif terbukti benar, yaitu seberapa besar kemungkinan *consequent* terjadi jika *antecedent* terjadi. *Confidence* dihitung sebagai rasio jumlah transaksi yang mengandung kedua item (*antecedent* dan *consequent*) terhadap jumlah transaksi yang mengandung *antecedent* saja. *Confidence* yang tinggi menunjukkan hubungan yang kuat antara *antecedent* dan *consequent*.
3. *Lift* digunakan untuk mengukur kekuatan asosiasi antara *antecedent* dan *consequent* dibandingkan dengan jika keduanya terjadi secara independen. *Lift* dihitung dengan membagi *confidence* dengan *support* dari *consequent*. Nilai *lift* lebih dari 1 menunjukkan bahwa *antecedent* dan *consequent* lebih sering terjadi bersama daripada jika keduanya independen, sedangkan nilai *lift* kurang dari 1 menunjukkan asosiasi negatif.

Dengan menggunakan metrik-metrik tersebut, *Association Rule Mining* dapat menghasilkan aturan-aturan yang paling relevan dan bermanfaat untuk mendukung pengambilan keputusan bisnis, seperti penataan produk, promosi, dan rekomendasi produk secara otomatis (Dwi Insani & Al Fatta, 2023).

2.1.7 Algoritma *Apriori*

Algoritma *Apriori* adalah salah satu algoritma utama dalam data mining yang digunakan untuk menemukan *frequent itemsets* dan membentuk aturan asosiasi dari data transaksi. Prinsip dasar algoritma ini adalah *anti-monotonicity*, yaitu sebuah itemset hanya dapat dikategorikan sebagai *frequent* jika semua *subset*-nya juga *frequent*. Dengan demikian, *Apriori* dapat secara efisien mengurangi jumlah kandidat itemset yang perlu diperiksa dalam proses pencarian pola.

Proses kerja algoritma *Apriori* dimulai dengan menentukan nilai minimum *support* dan *confidence*. Selanjutnya, algoritma membentuk kandidat itemset dari data transaksi, kemudian menghitung nilai *support* untuk setiap kandidat tersebut. Hanya *itemset* yang memenuhi nilai minimum *support* yang akan diproses ke tahap berikutnya. Proses ini dilakukan secara iteratif, mulai dari itemset tunggal (*1-itemset*), kemudian *2-itemset*, dan seterusnya, hingga tidak ada lagi kandidat itemset yang memenuhi syarat. Setelah *frequent itemset* ditemukan, algoritma membentuk aturan asosiasi yang memenuhi nilai minimum *confidence*, sehingga aturan yang dihasilkan benar-benar relevan dan memiliki probabilitas tinggi untuk terjadi secara bersamaan dalam transaksi.

Penelitian terbaru menunjukkan bahwa algoritma *Apriori* sangat efektif dalam mengidentifikasi pola pembelian konsumen dan dapat diterapkan pada

berbagai sektor, mulai dari ritel, farmasi, hingga kecantikan. Misalnya, penelitian oleh (Ashari et al., 2022) dan (Suhada et al., 2020) menunjukkan bahwa *Apriori* dapat menghasilkan pola asosiasi yang akurat, mendukung pengelolaan stok, serta meningkatkan strategi pemasaran dan penataan produk. Studi lain juga menyoroti penerapan *Apriori* dalam sistem *Point of Sale* (POS) berbasis web untuk memberikan rekomendasi produk dan meningkatkan efisiensi operasional toko.

Namun, algoritma *Apriori* memiliki kelemahan dalam hal efisiensi waktu dan penggunaan memori, terutama ketika diterapkan pada dataset yang sangat besar. Hal ini disebabkan oleh proses pembentukan kandidat *itemset* yang eksplisit dan perhitungan support yang berulang kali. Oleh karena itu, dalam praktiknya, pemilihan algoritma harus mempertimbangkan karakteristik data dan kebutuhan analisis.

2.1.8 Algoritma *FP-Growth*

FP-Growth adalah algoritma yang menggunakan struktur *FP-Tree* untuk menemukan *frequent itemsets* tanpa perlu membentuk kandidat. (Muhamad Andika et al., 2023) menyatakan bahwa *FP-Tree* mengompresi *database* transaksi ke dalam struktur pohon untuk mempercepat pencarian pola. Struktur *FP-Tree* terdiri dari:

1. *Node* yang merepresentasikan item dan jumlah kemunculannya.
1. *Header Table* yang menghubungkan *node-node* sejenis.
2. Jalur transaksi yang membentuk pola pembelian. Contohnya, transaksi {roti, susu, selai} dan {roti, selai} akan membentuk jalur pohon dengan *root* "roti" yang bercabang ke "susu" dan "selai".

Keunggulan FP-Growth adalah kemampuannya untuk menangani *dataset* besar dengan efisien karena tidak perlu membentuk kandidat *itemset* secara eksplisit, sehingga menghemat waktu dan memori. Penelitian terbaru juga menunjukkan bahwa *FP-Growth* lebih cepat dibandingkan *Apriori* dalam mengeksekusi analisis pola pembelian, tanpa mengurangi akurasi hasil aturan asosiasi yang dihasilkan.

FP-Growth banyak digunakan dalam *market basket analysis* untuk menemukan pola pembelian konsumen, mendukung strategi pemasaran, rekomendasi produk, dan pengelolaan inventori secara optimal.

2.1.9 Prediksi/Forecasting

Prediksi atau *forecasting* merupakan suatu aktivitas yang dilakukan untuk memperkirakan kejadian atau kebutuhan di masa depan dengan menggunakan pendekatan ilmiah dan analisis data (Ashari et al., 2022). prediksi adalah aktivitas yang dilakukan oleh peneliti untuk memperkirakan peristiwa yang akan terjadi di masa depan dengan menggunakan pendekatan ilmiah tertentu. Dalam konteks bisnis dan penelitian, *forecasting* sangat penting untuk mendukung pengambilan keputusan yang lebih tepat, terutama dalam hal perencanaan persediaan, penjualan, maupun strategi pemasaran.

Prediksi atau *forecasting* dapat didefinisikan sebagai upaya menentukan jumlah kebutuhan di bulan mendatang berdasarkan analisis data historis (*historical data*) atau serangkaian waktu tertentu. Data historis yang telah dikumpulkan dianalisis menggunakan metode statistik, matematis, atau algoritma tertentu untuk memperkirakan kebutuhan, permintaan, atau *tren* di masa

mendatang. Proses ini sangat membantu perusahaan dalam mengantisipasi perubahan pasar, mengelola stok, dan meningkatkan efisiensi operasional.

Penelitian terbaru juga menekankan pentingnya penerapan *forecasting* berbasis *data mining* dalam bisnis ritel untuk meningkatkan akurasi prediksi permintaan dan mengoptimalkan rantai pasok (Ustadatin et al., 2023).

Dengan demikian, prediksi atau *forecasting* menjadi salah satu komponen penting dalam sistem pengambilan keputusan berbasis data, khususnya dalam pengelolaan bisnis ritel yang dinamis.

2.1.10 Penjualan

Penjualan merupakan salah satu aktivitas utama dalam dunia bisnis, yaitu proses memindahkan barang atau jasa dari produsen atau penjual kepada konsumen dengan tujuan memperoleh keuntungan. Penjualan tidak hanya sekadar transaksi pertukaran barang atau jasa dengan uang, tetapi juga mencakup berbagai strategi dan upaya untuk memenuhi kebutuhan serta keinginan konsumen secara efektif.

Menurut (Tarigan et al., 2022), memahami pola penjualan sangat penting untuk meningkatkan profitabilitas bisnis. Dengan menganalisis data penjualan, perusahaan dapat mengetahui produk apa saja yang paling diminati, waktu-waktu penjualan tertinggi, serta perilaku konsumen dalam melakukan pembelian. Informasi ini sangat bermanfaat dalam merancang strategi pemasaran, pengelolaan stok, penentuan harga, serta pengembangan produk baru.

Penjualan yang efektif tidak hanya berfokus pada peningkatan volume transaksi, tetapi juga pada penciptaan nilai tambah bagi konsumen dan

membangun hubungan jangka panjang. Dalam *era digital* saat ini, analisis data penjualan menjadi semakin penting karena dapat memberikan *insight* yang mendalam melalui penerapan berbagai teknik analisis, seperti *market basket analysis*, *forecasting*, dan *data mining*.

Penelitian terbaru juga menegaskan bahwa pemanfaatan data penjualan secara optimal dapat membantu bisnis dalam mengidentifikasi peluang pasar, mengurangi risiko kerugian, serta meningkatkan daya saing di pasar yang semakin kompetitif (Sope, 2023).

Dengan demikian, penjualan bukan hanya aktivitas transaksi, tetapi juga proses strategis yang memerlukan pemahaman mendalam terhadap data dan perilaku konsumen untuk mencapai tujuan bisnis secara berkelanjutan.

2.2 Penelitian yang Relevan

Berikut adalah beberapa penelitian yang relevan dengan topik:

2.2.1 Comparison of Market Basket Analysis to Determine Consumer Purchasing Patterns Using FP-Growth and Apriori Algorithm

Penelitian ini berfokus pada perbandingan algoritma *Apriori* dan *FP-Growth* dalam analisis pola pembelian konsumen menggunakan data transaksi ritel. Tujuan utama dari penelitian ini adalah untuk mengevaluasi efektivitas kedua algoritma dalam menghasilkan aturan asosiasi yang dapat digunakan untuk mendukung pengambilan keputusan bisnis, seperti pengaturan tata letak produk di toko atau rekomendasi produk. Hasil penelitian menunjukkan bahwa *FP-Growth* lebih cepat dalam mengeksekusi data besar karena algoritma ini menghilangkan kebutuhan untuk menghasilkan kandidat *itemset* yang sering terjadi, yang dapat memperlambat *Apriori*. Sebaliknya, *Apriori* lebih efisien

ketika diterapkan pada dataset kecil, karena memerlukan lebih sedikit memori dan proses iteratif yang lebih sederhana. Penelitian ini menekankan bahwa pemilihan algoritma yang tepat sangat bergantung pada ukuran dataset yang dianalisis. Penelitian ini juga menunjukkan relevansinya dengan bisnis *bakery*, di mana penggunaan *FP-Growth* dapat membantu dalam mengidentifikasi kombinasi produk populer yang sering dibeli bersamaan, seperti roti dan kopi, terutama pada periode puncak penjualan (Aldino et al., 2021).

2.2.2 Market Basket Analysis Using Apriori and FP Growth Algorithm

Penelitian ini mengeksplorasi penggunaan algoritma *Apriori* dan *FP-Growth* dalam menganalisis data transaksi untuk menemukan hubungan antar produk yang sering dibeli bersama. Penelitian ini menggunakan *dataset* transaksi ritel dan menerapkan parameter seperti minimum *support* dan *confidence* untuk menentukan aturan asosiasi yang signifikan. *Apriori* menggunakan pendekatan iteratif untuk mencari pola-pola *itemset*, sementara *FP-Growth* menghindari iterasi dengan memanfaatkan struktur pohon untuk mengekstrak *itemset* yang sering muncul. Hasil penelitian menunjukkan bahwa *FP-Growth* lebih unggul dalam hal kecepatan karena tidak memerlukan pemeriksaan ulang *subset*, yang sering kali memperlambat *Apriori* pada dataset besar. Walaupun demikian, *Apriori* tetap relevan pada skenario dengan dataset kecil dan sederhana karena implementasinya yang lebih mudah dipahami dan digunakan. Penelitian ini memberikan wawasan yang berguna tentang bagaimana kedua algoritma dapat digunakan secara strategis dalam analisis *market basket* untuk membantu bisnis meningkatkan penjualan. Dalam konteks *bakery*, pemilihan algoritma yang tepat dapat meningkatkan efektivitas rekomendasi produk berdasarkan pola pembelian

pelanggan, misalnya, menemukan kombinasi roti dan minuman yang sering dibeli bersama (Hossain et al., 2019).

2.2.3 Evaluating The Performance Of Association Rules In Apriori And FP-Growth

Algorithms: Market Basket Analysis To Discoverrules Of Item Combinations

Penelitian ini mengkaji kinerja *Apriori* dan *FP-Growth* dalam menemukan aturan asosiasi untuk kombinasi item menggunakan data transaksi yang sangat besar dari PT. XYZ, sebuah perusahaan farmasi di Indonesia. Penelitian ini mengembangkan metrik baru seperti *Bi-confidence* dan *Bi-lift* untuk memberikan wawasan lebih dalam tentang hubungan antara *itemset*. Hasil analisis menunjukkan bahwa *FP-Growth* lebih cepat dalam eksekusi, sementara *Apriori* lebih efisien dalam penggunaan memori, khususnya ketika mengatasi *dataset* yang lebih kecil. Penelitian ini juga menekankan pentingnya memilih algoritma yang tepat untuk meningkatkan efisiensi bisnis dan *profitabilitas*, serta memberikan rekomendasi terkait strategi pemasaran dan manajemen inventaris berdasarkan pola pembelian konsumen. Dalam konteks *bakery*, temuan ini relevan karena mengidentifikasi produk-produk seperti roti dan kue yang sering dibeli bersama dan bisa digunakan untuk merancang strategi promosi lebih efektif berdasarkan pola konsumsi konsumen (Dwiputra et al., 2023).

2.2.4 Perbandingan Algoritma Apriori dan FP-Growth Terhadap Market Basket Analysis Pada Data Penjualan Toko Roti

Penelitian ini secara langsung membahas penerapan *Apriori* dan *FP-Growth* dalam menganalisis pola pembelian konsumen di industri *bakery*. Tujuan utama penelitian ini adalah untuk membandingkan kinerja kedua algoritma dalam menemukan asosiasi antar produk *bakery* yang sering dibeli bersama, berdasarkan

pola kombinasi dan nilai confidence. Hasil penelitian menunjukkan bahwa *FP-Growth* memiliki nilai *confidence* yang lebih tinggi dibandingkan *Apriori* pada kombinasi produk seperti pastry dan kopi, yang tercatat mencapai 55,21% untuk *FP-Growth*, sedangkan *Apriori* hanya mencatat 54,06%. Penelitian ini juga menunjukkan bahwa *FP-Growth* lebih efisien dalam mengolah data penjualan *bakery* yang cenderung lebih kompleks, dengan kombinasi produk yang lebih banyak. Oleh karena itu, *FP-Growth* lebih efektif dalam konteks *bakery* untuk mengidentifikasi pola-pola pembelian yang sering terjadi antara produk roti, kue, dan minuman (Fathurrahman et al., 2023).

2.2.5 Rancang Bangun Aplikasi Web untuk Meningkatkan Penjualan melalui Analisis Pola Pembelian

Penelitian ini bertujuan mengembangkan aplikasi berbasis *web* untuk menganalisis pola pembelian konsumen menggunakan algoritma *Apriori* dan *FP-Growth*. Aplikasi ini mengevaluasi performa kedua algoritma dalam menemukan pola transaksi yang dapat digunakan untuk meningkatkan penjualan melalui strategi promosi yang lebih tepat. Hasil penelitian menunjukkan bahwa *FP-Growth* lebih cepat dalam memproses data besar dan dapat memberikan informasi lebih tepat waktu kepada pemilik usaha. Dalam konteks *bakery*, aplikasi semacam ini dapat membantu pemilik usaha untuk merancang promosi berbasis data transaksi yang lebih dinamis, seperti diskon pada kombinasi produk roti dan minuman yang sering dibeli Bersama (NURSASONGKA, 2024).

2.2.6 Implementasi Algoritma *Apriori* pada Penjualan Alat Teknik Pertanian

Penelitian ini mengeksplorasi penerapan algoritma *Apriori* dalam analisis pola pembelian konsumen pada penjualan alat teknik pertanian. Meskipun

berbeda sektor, prinsip dari penelitian ini relevan dengan *market basket analysis* pada *bakery*, karena penelitian ini mengidentifikasi pola pembelian produk yang sering terjadi dalam kombinasi tertentu. Misalnya, jika pelanggan membeli Tali Kuralon, mereka cenderung juga membeli Pompa Submersible, dengan nilai *confidence* yang tinggi. Dalam konteks *bakery*, hal ini dapat diadaptasi untuk menemukan pola pembelian antara produk-produk yang sering dibeli bersamaan seperti roti dan selai, atau kue dan minuman. Penelitian ini menunjukkan pentingnya pemahaman tentang pola pembelian konsumen dalam mengoptimalkan persediaan produk dan strategi promosi di toko (Gunawan & Nataliani, 2021).

2.2.7 Penerapan Algoritma FP-Growth Untuk Menentukan Pola Pembelian Konsumen Pada AHASS Cibadak

Penelitian ini menggunakan FP-Growth untuk mengekstraksi pola pembelian pada transaksi suku cadang sepeda motor, dengan fokus pada pengoptimalan manajemen inventaris dan proses pemesanan. Meskipun sektor yang diteliti berbeda, prinsip yang digunakan—menganalisis pola pembelian produk yang sering dibeli bersama—dapat diaplikasikan pada *market basket analysis* dalam bisnis *bakery*. Misalnya, produk seperti roti atau kue dapat dianalisis bersama dengan produk lainnya, seperti minuman atau bahan-bahan tambahan, untuk mengidentifikasi kombinasi pembelian yang sering terjadi. Hasil penelitian ini menunjukkan bagaimana menggunakan FP-Growth untuk menemukan hubungan yang kuat antar produk, yang juga dapat meningkatkan kepuasan pelanggan dan mengoptimalkan strategi penempatan produk (Suhada et al., 2020).

2.2.8 Implementasi Algoritma *FP-Growth* Untuk Menemukan Pola Keterkaitan Antara Mata Kuliah Pemrograman Dan Mata kuliah Matematika

Penelitian ini menggunakan *FP-Growth* untuk menemukan pola keterkaitan antara mata kuliah pemrograman dan matematika di kalangan mahasiswa, dengan analisis terhadap data nilai mahasiswa. Walaupun konteks penelitian ini berbeda, konsep pola keterkaitan antara dua *item* dalam hal ini mata kuliah dapat diterapkan dalam *market basket analysis* untuk produk *bakery*. Misalnya, jika pelanggan membeli roti, mereka cenderung juga membeli selai, atau jika membeli kue, mereka juga membeli kopi. Penelitian ini menunjukkan bahwa *FP-Growth* dapat digunakan untuk menemukan pola-pola tak terduga yang memiliki tingkat *confidence* tinggi, yang berguna dalam merancang strategi promosi dan penempatan produk di toko bakery (Putri. P et al., 2021).

2.2.9 Implementasi *Market Basket Analysis* Pada Data Penjualan CV.XYZ Menggunakan Algoritma *FP-Growth*

Penelitian ini mengkaji penerapan *FP-Growth* dalam *market basket analysis* untuk data penjualan produk dari CV. XYZ. Hasil penelitian menunjukkan bahwa *FP-Growth* efektif dalam menghasilkan aturan asosiasi dan dapat membantu dalam pengelolaan inventaris serta strategi pemasaran. Misalnya, dengan mengidentifikasi produk yang sering dibeli bersama, strategi *bundling* dapat diterapkan untuk meningkatkan penjualan. Dalam konteks *bakery*, penelitian ini memberikan wawasan bagaimana *FP-Growth* dapat digunakan untuk mengidentifikasi kombinasi produk *bakery* yang paling sering dibeli bersama, seperti roti dan kopi, untuk membantu merancang penawaran paket atau promosi yang lebih tepat sasaran (Ridho et al., 2024).

2.2.10 Market Basket Analysis Of Administrative Patterns Data Of Consumer Purchases Using Data Mining Technology

Penelitian ini membahas penggunaan algoritma *Apriori* dalam *market basket analysis* untuk menganalisis pola pembelian konsumen di supermarket. Penelitian ini menunjukkan pentingnya pemahaman tentang kombinasi produk yang sering dibeli bersamaan untuk meningkatkan manajemen inventaris dan strategi pemasaran. Meskipun fokus pada *supermarket*, prinsip yang ditemukan dalam penelitian ini—menganalisis asosiasi antar *item* dan memanfaatkan nilai *support* dan *confidence*—dapat diterapkan dalam *market basket analysis* di *bakery*. Dengan menggunakan pendekatan ini, pemilik *bakery* dapat mengidentifikasi produk mana yang sering dibeli bersamaan, seperti kue dan minuman, untuk menyusun strategi promosi yang lebih efektif, seperti menawarkan diskon atau *bundling* produk (Samboteng et al., 2022).

2.2.11 Application of Market Basket Analysis on Beauty Clinic to Increasing Customer's Buying Decision

Penelitian ini mengeksplorasi bagaimana *market basket analysis* dapat digunakan untuk meningkatkan keputusan pembelian pelanggan di klinik kecantikan dengan menggunakan FP-Growth. Dengan mengidentifikasi pola pembelian pelanggan menggunakan FP-Growth, klinik dapat merancang paket promosi yang lebih baik. Walaupun sektor yang diteliti berbeda, pendekatan yang sama dapat diterapkan di bisnis *bakery*. Misalnya, dengan menemukan kombinasi produk *bakery* yang sering dibeli bersama, seperti roti dan selai, pemilik *bakery* dapat merancang promosi paket yang menarik, atau strategi penempatan produk yang lebih optimal. Penelitian ini menunjukkan bahwa menggunakan data transaksi yang dianalisis dengan FP-Growth dapat memberikan wawasan yang

sangat berharga dalam merancang strategi pemasaran yang lebih efektif dan menarik bagi pelanggan (Dio et al., 2023).

2.2.12 Implementasi *Data Mining* Menggunakan Algoritma *Apriori* Dalam Menentukan Persediaan Barang (Studi Kasus: Toko Sinar Harahap)

Penelitian ini membahas penerapan *Apriori* dalam data mining untuk menganalisis transaksi penjualan barang di Toko Sinar Harahap, dengan tujuan untuk menentukan produk yang memiliki permintaan tinggi. Penelitian ini mengidentifikasi pola asosiasi yang kuat, seperti adanya hubungan antara beras dan telur, dengan *confidence* sebesar 62%, yang menunjukkan bahwa pembelian beras sering diikuti dengan pembelian telur. Dalam konteks *bakery*, pola serupa dapat ditemukan dengan menganalisis produk seperti roti dan selai, yang sering dibeli bersama oleh konsumen. Penelitian ini menunjukkan bahwa *Apriori* dapat membantu dalam mengidentifikasi kombinasi produk yang sering dibeli dan dapat digunakan untuk merencanakan pengadaan stok serta strategi promosi yang lebih efisien di toko bakery (Tarigan et al., 2022).

2.2.13 Implementasi Metode *Market Basket Analysis* Pada Sistem Penjualan Dan Pemasaran Butik Nazwa Fashion

Penelitian ini membahas penerapan *Market Basket Analysis* (MBA) pada sistem penjualan dan pemasaran butik *online*, dengan menggunakan algoritma untuk memahami kombinasi barang yang sering dibeli bersama. Meskipun fokusnya pada butik *fashion*, prinsip yang diterapkan—menemukan kombinasi produk yang sering dibeli bersamaan—dapat diterapkan dalam *market basket analysis* untuk produk *bakery*. Dengan mengetahui produk mana yang sering dibeli bersamaan, seperti roti dan minuman, pemilik toko bakery dapat merancang

strategi *bundling* atau diskon untuk produk-produk tersebut. Penelitian ini menunjukkan bagaimana MBA dapat membantu dalam merancang sistem penjualan *online* yang lebih efektif, serta meningkatkan pengalaman belanja pelanggan dengan memberikan rekomendasi produk yang relevan (Wati et al., 2020).

2.2.14 Market Basket Analysis Apotek Berbasis Web Menggunakan Metode Algoritma Apriori

Penelitian ini membahas penerapan *Apriori* dalam *market basket analysis* untuk menganalisis pola pembelian di apotek. Penelitian ini menunjukkan bagaimana *Apriori* dapat membantu apotek dalam memahami pola pembelian obat-obatan, serta mengoptimalkan penempatan produk dan persediaan berdasarkan hasil analisis. Walaupun konteksnya apotek, prinsip yang digunakan dapat diterapkan dalam bisnis *bakery*. Sebagai contoh, jika produk seperti roti atau kue sering dibeli bersamaan dengan produk lainnya, seperti minuman atau selai, maka *Apriori* dapat digunakan untuk merancang strategi penempatan produk atau promosi yang lebih terfokus (Bashiir & Witanti, 2024).

2.2.15 Comparison of Apriori, Apriori-TID, and FP-Growth Algorithms in Market Basket Analysis at Grocery Stores

Penelitian ini mengkaji perbandingan tiga algoritma yang digunakan dalam analisis keranjang belanja di toko grosir, yaitu Apriori, Apriori-TID, dan FP-Growth. Evaluasi dilakukan berdasarkan waktu eksekusi, jumlah aturan asosiasi yang dihasilkan, dan penggunaan memori. Hasil eksperimen menunjukkan bahwa FP-Growth memiliki waktu eksekusi yang paling cepat, terutama pada dataset besar, meskipun memerlukan memori yang sedikit lebih

banyak dibandingkan *Apriori* dan *Apriori-TID*. *Apriori-TID* cenderung lebih efektif pada dataset kecil hingga menengah, dengan keunggulan dalam hal penggunaan memori. Penelitian ini menyoroti bahwa dalam konteks *bakery*, penggunaan *FP-Growth* sangat disarankan untuk analisis data transaksi besar dan kompleks, seperti saat ada promosi musiman, karena lebih efisien dalam mengekstrak aturan asosiasi yang banyak dan relevan .



BAB III

METODOLOGI PENELITIAN

3.1 Analisa Kebutuhan

Analisa kebutuhan dilakukan untuk mengidentifikasi seluruh komponen yang diperlukan dalam mendukung proses penelitian, mulai dari sumber data, perangkat lunak yang digunakan, hingga pustaka pemrograman yang relevan. Penelitian ini berfokus pada penerapan algoritma *Apriori* dan *FP-Growth* dalam menganalisis pola pembelian konsumen melalui pendekatan *market basket analysis*, sehingga kebutuhan sistem yang dirancang harus mampu menunjang proses *frequent itemset mining* dan evaluasi aturan asosiasi secara efisien dan sistematis.

3.1.1 Kebutuhan Data

Data yang digunakan dalam penelitian ini merupakan data transaksi penjualan produk pada sebuah toko bakery, diperoleh dari platform *Kaggle* dan dapat diakses melalui tautan <https://www.kaggle.com/datasets/akashdeepkuila/bakery>. Dataset Bakery.csv disusun oleh Akashdeep Kuila pada tahun 2022 dan terdiri atas 18.887 baris data transaksi.

	A	B	C	D	E
1	TransactionNo	Items	DateTime	Daypart	DayType
2	1	Bread	10/30/2016 9:58	Morning	Weekend
3	2	Scandinavian	10/30/2016 10:05	Morning	Weekend
4	2	Scandinavian	10/30/2016 10:05	Morning	Weekend
5	3	Hot chocolate	10/30/2016 10:07	Morning	Weekend
6	3	Jam	10/30/2016 10:07	Morning	Weekend
7	3	Cookies	10/30/2016 10:07	Morning	Weekend
8	4	Muffin	10/30/2016 10:08	Morning	Weekend
9	5	Coffee	10/30/2016 10:13	Morning	Weekend
10	5	Pastry	10/30/2016 10:13	Morning	Weekend
11	5	Bread	10/30/2016 10:13	Morning	Weekend
12	6	Medialuna	10/30/2016 10:16	Morning	Weekend
13	6	Pastry	10/30/2016 10:16	Morning	Weekend
14	6	Muffin	10/30/2016 10:16	Morning	Weekend
15	7	Medialuna	10/30/2016 10:19	Morning	Weekend
16	7	Pastry	10/30/2016 10:19	Morning	Weekend
17	7	Coffee	10/30/2016 10:19	Morning	Weekend
18	7	Tea	10/30/2016 10:19	Morning	Weekend
19	8	Pastry	10/30/2016 10:20	Morning	Weekend
20	8	Bread	10/30/2016 10:20	Morning	Weekend
21	9	Bread	10/30/2016 10:21	Morning	Weekend
22	9	Muffin	10/30/2016 10:21	Morning	Weekend
23	10	Scandinavian	10/30/2016 10:25	Morning	Weekend
24	10	Medialuna	10/30/2016 10:25	Morning	Weekend

Gambar 3.1 Gambar *Dataset Asli*

Setiap baris pada *dataset* ini merepresentasikan satu *produk* yang dibeli oleh pelanggan dalam suatu transaksi. Beberapa baris dengan *TransactionNo* yang sama menunjukkan bahwa *item* tersebut dibeli secara bersamaan dalam satu transaksi. Secara struktural, dataset ini terdiri atas lima atribut utama sebagaimana disajikan dalam tabel berikut:

Tabel 3.1 Keterangan Atribut *Dataset Bakery*

Atribut	Keterangan	Variabel
TransactionNo	Nomor unik untuk mengidentifikasi setiap transaksi.	Integer
Items	Nama barang yang dibeli dalam transaksi.	String
DateTime	Waktu dan tanggal saat transaksi terjadi.	Datetime
Daypart	Segmentasi waktu (Morning, Afternoon, Evening, Night).	String
DayType	Jenis hari (Weekday atau Weekend).	String

Untuk kepentingan analisis asosiasi, penelitian ini hanya menggunakan tiga atribut, yaitu *TransactionNo*, *Items*, dan *DateTime*, karena ketiganya relevan dalam membentuk struktur transaksi dan melakukan ekstraksi *itemset*. Atribut lain seperti *Daypart* dan *DayType* tidak digunakan dalam proses *mining* karena tidak memiliki pengaruh langsung terhadap pembentukan aturan asosiasi antar item.

Setelah memahami struktur *dataset*, dilakukan analisis deskriptif untuk mengevaluasi karakteristik data dan mendukung pemahaman awal terhadap pola transaksi pelanggan. Hasil eksplorasi ditampilkan dan dijelaskan sebagai berikut:

1. Jumlah Item Unik

Dataset ini memuat sebanyak 94 produk unik yang meliputi berbagai jenis minuman (*seperti Coffee, Tea, Hot Chocolate*), makanan ringan (*Pastry, Cake*), serta produk utama seperti Bread. Keberagaman ini memberikan ruang yang luas untuk eksplorasi pola asosiasi antar produk.

2. Jumlah Transaksi dan Ukuran Transaksi

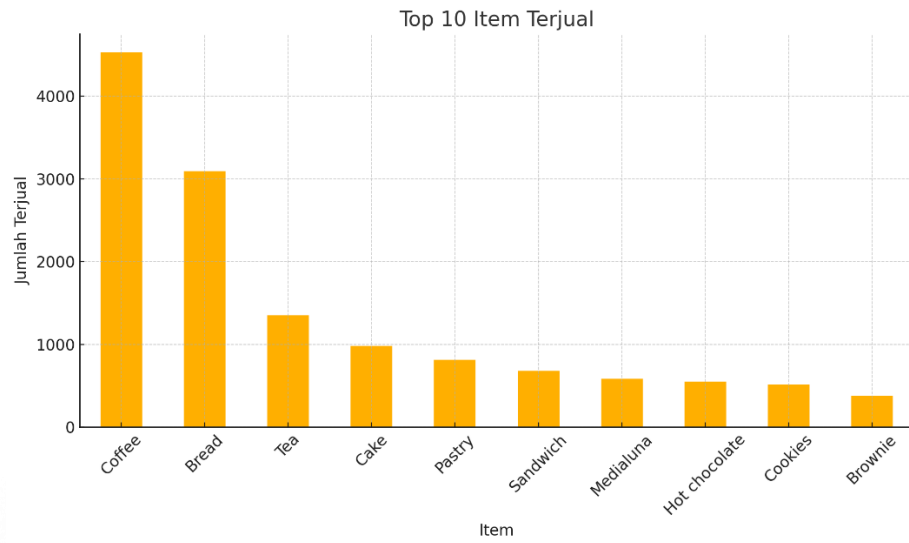
Dari total 18.887 baris data, ditemukan sekitar 9.465 transaksi unik berdasarkan *TransactionNo*. Rata-rata jumlah *item* per transaksi adalah 2 item. Sebagian besar transaksi (sekitar 40%) hanya terdiri dari satu *item*, sedangkan transaksi yang mencakup tiga atau lebih item menurun secara *eksponensial*. Transaksi dengan jumlah item terbanyak mencapai hingga 10 item dalam satu pembelian.

3. Distribusi Waktu Transaksi

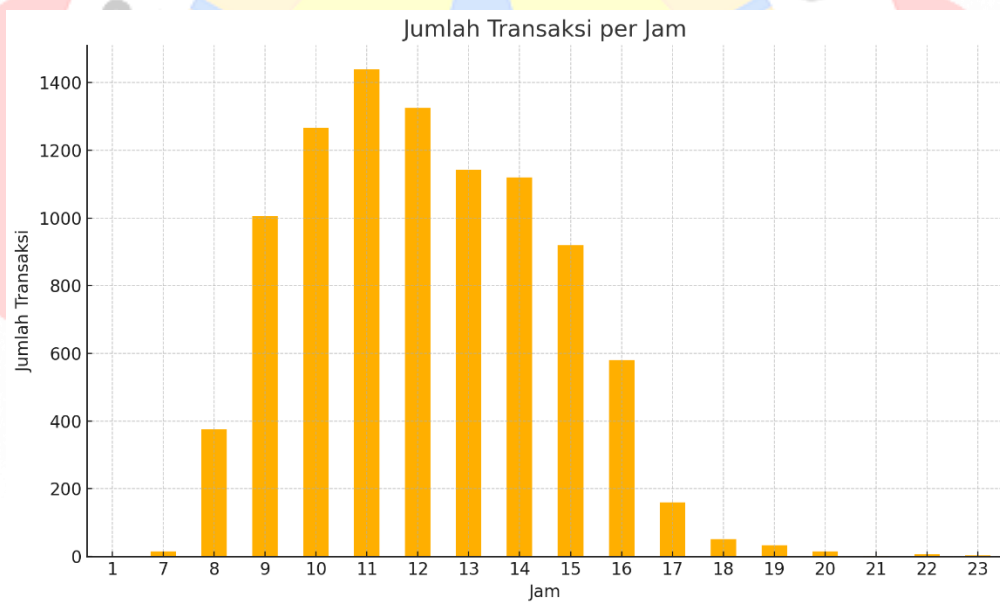
Berdasarkan analisis temporal terhadap kolom *DateTime*, ditemukan bahwa puncak aktivitas transaksi terjadi pada rentang pukul 09.00 hingga 15.00, dengan puncak tertinggi pada pukul 11.00–12.00 WIB. Hal ini menunjukkan bahwa waktu sarapan hingga makan siang merupakan waktu belanja dominan bagi pelanggan toko roti.

Secara distribusi harian, volume transaksi tertinggi terjadi pada hari Sabtu dan Minggu, mengindikasikan bahwa aktivitas pembelian meningkat signifikan saat akhir pekan. Hari dengan jumlah transaksi terendah adalah Rabu, yang menunjukkan pola penurunan aktivitas di pertengahan minggu.

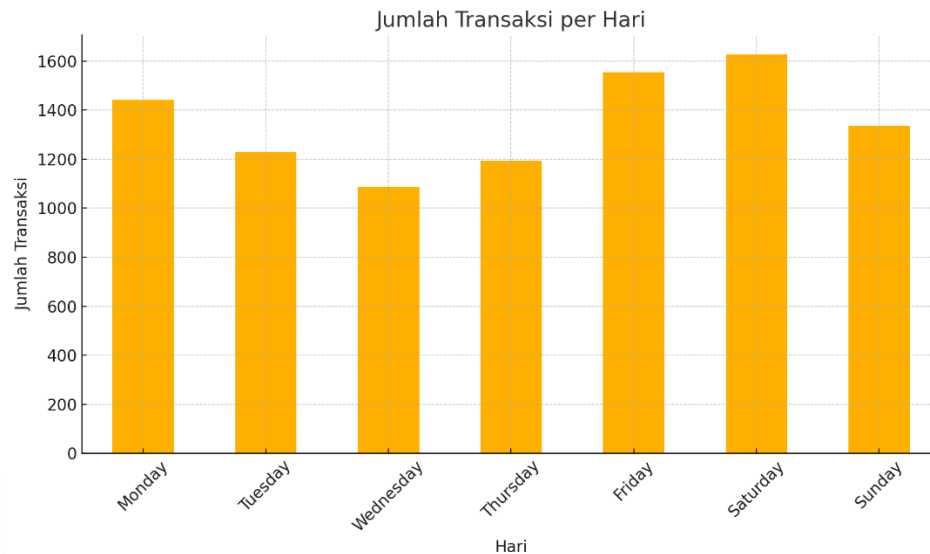
Untuk memberikan pemahaman yang lebih jelas, berikut visualisasi dari hasil eksplorasi data:



Gambar 3.2 Top 10 *Item* Terlaris



Gambar 3.3 Jumlah Transaksi per Jam



Gambar 3.4 Jumlah Transaksi per Hari

Visualisasi ini menunjukkan bahwa produk seperti *Coffee*, *Bread*, dan *Tea* adalah produk dengan volume penjualan tertinggi, menjadikannya kandidat utama untuk analisis asosiasi. Temuan ini juga memperkuat argumentasi untuk menerapkan *time-based targeting* dalam strategi promosi produk.

3.1.2 Kebutuhan *Tools* dan *Software*

Untuk mendukung eksperimen dan analisis data dalam penelitian ini, sejumlah perangkat lunak dan pustaka pemrograman yang relevan telah digunakan. Pemilihan *tools* ini mempertimbangkan fleksibilitas, kemudahan integrasi, serta kemampuannya dalam mendukung pemrograman *Python* untuk menerapkan algoritma *Apriori* dan *FP-Growth*. Rincian kebutuhan *tools* disajikan pada tabel berikut:

Tabel 3.2 Kebutuhan *Tools* dan *Software*

No.	Tools/Software	Fungsi
1	Python	Bahasa pemrograman utama untuk analisis data dan penerapan algoritma.
2	Visual Studio Code	Editor kode utama yang digunakan untuk menulis dan menjalankan skrip Python.
3	Streamlit	Framework untuk membangun antarmuka web interaktif yang memungkinkan visualisasi dan interaksi dengan hasil analisis secara langsung.
4	Replit	Platform cloud-based yang digunakan untuk menjalankan dan berbagi kode Python secara online.
5	Pyproject.toml	File konfigurasi proyek untuk mengelola dependensi dan pengaturan proyek Python.

Penggunaan *Visual Studio Code* sebagai *editor kode* utama memungkinkan pengembangan yang fleksibel dan efisien. *Streamlit* digunakan untuk membangun aplikasi berbasis *web* interaktif, mempermudah visualisasi hasil analisis dan interaksi pengguna dengan data. *Platform Replit* digunakan untuk menjalankan kode *Python* secara *cloud-based*, sehingga dapat diakses dari berbagai perangkat dan mendukung eksperimen secara kolaboratif.

3.1.3 Kebutuhan *Library*

Berikut adalah *pustaka pemrograman* yang digunakan untuk mendukung proses *preprocessing* data, penerapan algoritma *Apriori* dan *FP-Growth*, serta visualisasi hasil analisis:

Tabel 3.3 *Library* yang Digunakan

No.	Library	Bahasa	Fungsi
1	Pandas	Python	Untuk manipulasi, filter, dan pembersihan data dalam format tabular.
2	Mlxtend	Python	Mengimplementasikan algoritma <i>Apriori</i> dan <i>FP-Growth</i> secara efisien.
3	Matplotlib, Seaborn	Python	Digunakan untuk menyajikan visualisasi data seperti grafik frekuensi dan distribusi.
4	Streamlit	Python	Digunakan untuk membangun antarmuka web interaktif yang menampilkan hasil analisis secara langsung.

Pustaka *Mlxtend* dipilih karena kemampuannya dalam menangani algoritma *Apriori* dan *FP-Growth* dengan efisien dalam konteks data transaksi. *Pandas* digunakan untuk manipulasi data sebelum analisis dimulai, sementara *Matplotlib* memberikan alat untuk visualisasi hasil yang mudah dipahami. *Streamlit* memastikan bahwa hasil analisis dapat dipresentasikan dalam bentuk *web* yang interaktif, memungkinkan pengguna untuk mengeksplorasi hasil eksperimen secara langsung.

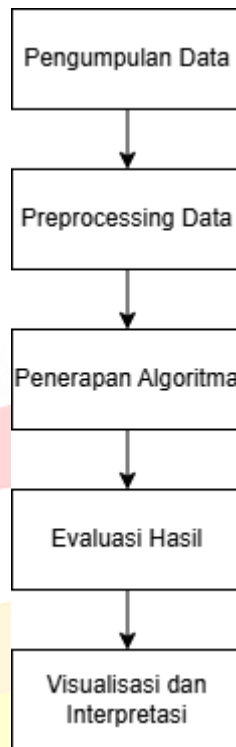
3.2 Pembahasan Metode yang Digunakan

Penelitian ini menerapkan metode *Association Rule Mining (ARM)*, yaitu salah satu pendekatan analitis dalam *data mining* yang bertujuan untuk mengidentifikasi hubungan atau keterkaitan antar item dalam suatu himpunan data transaksi. ARM digunakan secara luas dalam analisis pola perilaku konsumen, terutama dalam konteks *market basket analysis (MBA)*, untuk menemukan kombinasi produk yang sering dibeli secara bersamaan dalam satu transaksi.

Association Rule Mining sangat tepat diterapkan dalam konteks *market basket analysis*, yaitu teknik analisis yang digunakan untuk mengetahui kebiasaan belanja konsumen melalui pemetaan produk yang sering dibeli bersamaan. Dengan menerapkan metode ini, bisnis dapat memperoleh wawasan yang berguna untuk penempatan produk, promosi silang (*cross-selling*), dan rekomendasi produk.

3.2.1 Tahapan Metode Penelitian

Untuk mengimplementasikan metode *Association Rule Mining* dalam penelitian ini, dilakukan sejumlah tahapan sebagai berikut:



Gambar 3.5 Tahapan Metode Penelitian

1. Pengumpulan Data

Dataset yang digunakan merupakan data transaksi pembelian pada toko *bakery*, yang bersumber dari *Kaggle*. Dataset berisi 18.887 transaksi dengan berbagai jenis produk *bakery* yang dibeli oleh pelanggan.

2. *Preprocessing Data*

Tahap ini mencakup pembersihan data, penghapusan duplikasi, serta transformasi data ke dalam format transaksi. Atribut yang tidak relevan seperti *Daypart* dan *DayType* dihapus karena tidak memberikan kontribusi langsung terhadap pola pembelian.

3. Penerapan Algoritma *Apriori* dan *FP-Growth*

Setelah data siap, dilakukan proses mining dengan dua algoritma: *Apriori* dan *FP-Growth*. Kedua algoritma ini akan diterapkan secara terpisah untuk

menemukan aturan asosiasi berdasarkan nilai *minimum support* dan *minimum confidence* yang telah ditentukan. Hasil dari kedua algoritma akan dibandingkan untuk menilai perbedaan kinerjanya.

4. Evaluasi Hasil Mining

Hasil aturan asosiasi yang terbentuk akan dianalisis dan dievaluasi menggunakan beberapa metrik seperti jumlah aturan yang dihasilkan, waktu eksekusi, serta kekuatan hubungan antar item berdasarkan nilai *support*, *confidence*, dan *lift*.

5. Visualisasi dan Interpretasi Hasil

Untuk mempermudah pemahaman, hasil analisis ditampilkan dalam bentuk tabel dan grafik. Visualisasi dilakukan untuk menunjukkan frekuensi item yang paling sering dibeli serta pola waktu transaksi (jam dan hari).

3.2.2 Penentuan Parameter Minimum Support dan Minimum Confidence

Dalam *association rule mining*, dua parameter utama yang digunakan untuk menentukan kualitas aturan asosiasi yang ditemukan adalah *minimum support* dan *minimum confidence*. *Minimum support* mengukur frekuensi kemunculan suatu *itemset* dalam *dataset* transaksi, sedangkan *minimum confidence* mengukur kekuatan asosiasi antara item dalam aturan. Kombinasi kedua parameter ini mempengaruhi jumlah aturan yang dihasilkan serta kualitas aturan yang ditemukan dalam analisis *market basket*. Parameter yang terlalu ketat (*high support dan confidence*) akan menghasilkan sedikit aturan yang sangat kuat, sementara parameter yang terlalu longgar dapat menghasilkan banyak aturan yang kurang relevan (Wijayanto, H., & Purnomo, 2022). Oleh karena itu, pemilihan nilai untuk kedua parameter ini sangat penting untuk mencapai keseimbangan yang optimal antara jumlah aturan dan kualitas asosiasi yang dihasilkan.

Untuk eksperimen ini, dilakukan uji coba dengan beberapa kombinasi nilai minimum *support* dan minimum *confidence* untuk mengevaluasi dampaknya terhadap jumlah aturan yang dihasilkan, waktu eksekusi, serta kualitas asosiasi yang ditemukan. Nilai minimum *support* yang diuji adalah 0,01, 0,02, dan 0,05, sementara nilai minimum *confidence* yang diuji adalah 0,5, 0,6, dan 0,7. Kombinasi-kombinasi ini dipilih untuk mengeksplorasi pengaruh variasi dari nilai *support* dan *confidence* terhadap hasil eksperimen, dengan mempertimbangkan karakteristik dari *dataset* yang digunakan, yaitu data transaksi penjualan produk *bakery*.

Alasan Pemilihan Kombinasi Parameter Berdasarkan *Dataset Bakery.csv*:

1. Minimum *Support*:

- 1) *Support* 0,01: Dengan *dataset* yang berisi 18.887 transaksi dan 94 produk unik, nilai *support* 0,01 memungkinkan eksperimen untuk menangkap pola pembelian yang lebih jarang namun masih relevan secara bisnis. Pola pembelian ini bisa memberikan wawasan penting untuk produk-produk tertentu yang mungkin jarang dibeli bersama, tetapi relevan untuk promosi silang (Suryadi & Santosa, 2021).
- 2) *Support* 0,02 dan *Support* 0,05: Nilai-nilai *support* ini akan lebih selektif dalam menemukan aturan, sehingga memberikan asosiasi yang lebih kuat dan lebih umum, seperti produk yang sering dibeli bersama seperti *Coffee* dan *Pastry*, yang memberikan informasi lebih relevan untuk strategi pemasaran (Wijayanto & Purnomo, 2022).

2. Minimum Confidence:

1) *Confidence* 0,5: Nilai ini memungkinkan eksperimen untuk menemukan aturan dengan kepercayaan lebih rendah, yang berguna untuk menemukan pola asosiasi yang kurang kuat tetapi tetap relevan. Hal ini penting untuk strategi cross-selling atau promosi produk yang dapat menarik pembeli meskipun asosiasinya tidak terlalu kuat (Wijayanto & Purnomo, 2022).

2) *Confidence* 0,6 dan 0,7: Nilai confidence ini memastikan bahwa aturan yang ditemukan lebih kuat dan dapat diandalkan. Dengan *confidence* yang lebih tinggi, aturan yang ditemukan memiliki tingkat kepercayaan yang lebih tinggi, sehingga dapat digunakan untuk menyarankan produk dengan lebih tepat, terutama untuk produk yang sering dibeli bersama dalam jumlah besar (Suryadi & Santosa, 2021).

Kombinasi nilai minimum *support* dan minimum *confidence* yang digunakan dalam eksperimen ini dipilih berdasarkan penelitian sebelumnya dalam *market basket analysis* yang menunjukkan bahwa variasi dari kedua parameter ini memiliki dampak signifikan terhadap jumlah aturan dan kualitas asosiasi yang dihasilkan. Penelitian sebelumnya menunjukkan bahwa eksperimen dengan nilai *support* yang lebih rendah memungkinkan penemuan pola yang lebih jarang namun relevan, sementara nilai *confidence* yang lebih tinggi menghasilkan aturan yang lebih kuat dan dapat diandalkan (Suryadi, D., & Santosa, 2021).

3.2.4 Output yang Diharapkan

Setelah eksperimen dengan beberapa kombinasi parameter, *output* yang diharapkan dari penerapan metode *association rule mining* pada dataset *Bakery.csv* adalah sebagai berikut:

1. Aturan Asosiasi Antar Produk: Hasil yang diharapkan adalah terbentuknya aturan asosiasi yang menunjukkan hubungan antar produk yang sering dibeli bersama. Misalnya, aturan yang dihasilkan seperti $\{Coffee\} \rightarrow \{Pastry\}$, diharapkan dapat memberikan wawasan terhadap pola pembelian yang dominan dalam transaksi. Pola-pola ini berguna bagi pemilik usaha untuk merancang strategi pemasaran, promosi silang (*cross-selling*), hingga rekomendasi produk.
2. *Frequent Itemset* yang Relevan: Proses mining akan menghasilkan himpunan item (*itemset*) yang sering muncul dalam satu transaksi. Dengan menentukan nilai minimum *support* dan minimum *confidence*, hanya *itemset* yang memenuhi ambang batas tersebut yang akan dipertahankan sebagai frequent itemset. Hasil ini membantu dalam penyusunan paket produk dan bundling yang lebih efektif.
3. Perbandingan Kinerja Algoritma *Apriori* dan *FP-Growth*: Hasil eksperimen bertujuan untuk mengevaluasi dan membandingkan dua algoritma berdasarkan kinerja mereka dalam hal:
 - 1) Waktu Eksekusi: Mengukur kecepatan algoritma dalam menemukan aturan asosiasi pada dataset besar.
 - 2) Jumlah Aturan yang Dihasilkan: Membandingkan jumlah aturan yang dihasilkan dengan parameter yang berbeda dan memahami *trade-off* antara jumlah aturan dan kualitas aturan.

- 3) Kualitas Aturan: Menganalisis aturan berdasarkan *support*, *confidence*, dan *lift*. Aturan dengan nilai *confidence* dan *lift* tinggi dianggap lebih kuat dan relevan untuk diterapkan dalam konteks bisnis.
4. Visualisasi Pola Pembelian Konsumen: Sebagai bagian dari interpretasi hasil, pola-pola pembelian konsumen akan divisualisasikan dalam bentuk grafik batang, *pie chart*, atau heatmap untuk memudahkan pembacaan dan komunikasi hasil kepada pihak non-teknis, seperti pemilik usaha atau manajer pemasaran.

Dengan luaran yang jelas, penelitian ini diharapkan dapat memberikan kontribusi praktis bagi pelaku usaha kecil menengah (UKM) di bidang makanan dan minuman, serta kontribusi ilmiah dalam membandingkan efektivitas dua algoritma populer dalam *market basket analysis*.

3.3 Pembahasan Algoritma yang Digunakan

Dalam penelitian ini, digunakan dua algoritma utama dalam proses *association rule mining*, yaitu algoritma *Apriori* dan *FP-Growth*. Keduanya merupakan metode populer yang digunakan untuk mengekstraksi aturan asosiasi dari dataset transaksi. Meskipun keduanya memiliki tujuan yang sama, pendekatan dan efisiensi algoritma tersebut berbeda secara signifikan, sehingga perlu dianalisis secara terpisah.

3.3.1 Algoritma *Apriori*

Algoritma *Apriori* merupakan salah satu metode paling awal dan populer dalam analisis *market basket*, yang bertujuan menemukan pola asosiasi antara item dalam suatu kumpulan transaksi. Prinsip utama algoritma ini adalah bahwa suatu *itemset* hanya akan *frequent* jika semua *subset*-nya juga *frequent* (*Apriori*

property). *Apriori* bekerja dengan pendekatan *level-wise iteration*, yaitu memulai dari *frequent 1-itemset*, lalu mengembangkan ke *frequent k-itemset* secara iteratif dengan melakukan pruning berdasarkan minimum *support*. Langkah-langkah utama algoritma *Apriori* adalah sebagai berikut:

1. Melakukan pemindaian (*scan*) terhadap seluruh dataset untuk menghitung *support* masing-masing item (*1-itemset*).
2. Menghapus *item* yang memiliki *support* di bawah batas minimum (*pruning*).
3. Membentuk kombinasi *itemset* berikutnya (*2-itemset*, *3-itemset*, dst.) dari *frequent itemset* sebelumnya.
4. Mengulangi proses hingga tidak ada lagi kombinasi *itemset* yang memenuhi nilai minimum *support*.
5. Menghitung nilai *confidence* untuk membentuk aturan asosiasi dari *itemset* yang lolos.

Rumus Perhitungan dalam algoritma *Apriori*:

- a. *Support* (S) mengukur seberapa sering *itemset* muncul dalam dataset, dihitung sebagai:

$$Support(X) = \frac{\text{Jumlah Transaksi yg Mengandung itemset } X}{\text{Total Jumlah Transaksi}}$$

- b. *Confidence* (C) mengukur kekuatan hubungan antar *itemset*, yaitu seberapa besar kemungkinan *itemset consequent* terjadi jika *antecedent* terjadi, dihitung dengan:

$$Confidence(A \rightarrow B) = \frac{\text{Jumlah transaksi yang mengandung } A \text{ dan } B}{\text{Jumlah transaksi yg mengandung } A}$$

- c. *Lift* (L) mengukur kekuatan asosiasi antar *itemset* dibandingkan dengan jika keduanya terjadi secara independen, dihitung dengan:

$$Lift(A \rightarrow B) = \frac{Confidence(A \rightarrow B)}{Support(B)}$$

Rumus-rumus ini digunakan untuk mengevaluasi kekuatan dan kualitas dari aturan asosiasi yang dihasilkan. Aturan dengan nilai *confidence* dan *lift* tinggi dianggap lebih kuat dan relevan untuk diterapkan.

a. Ilustrasi Perhitungan Manual *Apriori* dengan Data Sampel

Untuk memberikan pemahaman konkret mengenai mekanisme algoritma *Apriori*, digunakan dataset sampel yang terdiri dari 10 transaksi sebagai berikut:

Tabel 3.4 Data sampel *apriori*

TransaktionNo	Item
1	Bread
2	Scandinavian
3	Hot Chocolate, Jam, Cookies
4	Muffin
5	Coffee, Pastry, Bread
6	Medialuna, Pastry, Muffin
7	Medialuna, Pastry, Coffee, Tea
8	Pastry, Bread
9	Bread, Muffin
10	Scandinavian, Medialuna

Parameter:

Minimum *Support*: 20% (setara dengan 2 transaksi)

Minimum *Confidence*: 60%

Tahap 1: *Frequent 1-Itemset (L1)*

Menghitung support masing-masing item:

Tabel 3.5 Menghitung *Support Apriori*

Item	Count	Support (%)	Lolos Min. Support
Bread	4	40%	Ya
Scandinavian	2	20%	Ya
Hot Chocolate	1	10%	Tidak
Jam	1	10%	Tidak
Cookies	1	10%	Tidak
Muffin	3	30%	Ya
Coffee	2	20%	Ya
Pastry	4	40%	Ya
Medialuna	3	30%	Ya
Tea	1	10%	Tidak

Tahap 2: *Frequent 2-Itemset (L2)*

Beberapa *itemset* yang memenuhi minimum *support*:

Tabel 3.6 *Itemset* yg memenuhi *support*

Itemset	Count	Support (%)	Lolos
Bread, Pastry	2	20%	Ya
Coffee, Pastry	2	20%	Ya
Pastry, Medialuna	2	20%	Ya

Tahap 3: Pembentukan Aturan Asosiasi

Tabel 3.7 Pembentukan Aturan Asosiasi

Rule	Support	Confidence (%)	Lolos Min. Confidence
Coffee → Pastry	2	100%	Ya
Medialuna → Pastry	2	66.67%	Ya
Bread → Pastry	2	50%	Tidak

- b. Implementasi dan Analisis Algoritma *Apriori* dengan *Dataset bakery.csv* (18.888 Transaksi)

Pada tahap ini, algoritma *Apriori* diimplementasikan untuk mengekstraksi aturan asosiasi dari dataset penjualan roti yang terdiri atas

9.465 transaksi unik dan 94 item berbeda, sebagaimana telah dijelaskan pada bagian analisis kebutuhan data.

Implementasi dilakukan menggunakan bahasa pemrograman *Python* dengan pustaka *mlxtend*, dan eksperimen dijalankan dengan nilai parameter minimum *support* sebesar 0.02 dan minimum *confidence* sebesar 0.5. Parameter ini dipilih untuk memastikan bahwa hanya *itemset* yang cukup sering dan memiliki tingkat asosiasi kuat yang akan dipertimbangkan dalam proses pembentukan aturan.

Berikut adalah semua aturan yang ditemukan menggunakan algoritma *Apriori*:

```
=== APRIORI RESULTS ===
Jumlah aturan      : 8
Waktu eksekusi     : 0.0227 detik
Penggunaan memori  : 0.74 MB

Semua aturan yang ditemukan:
```

	antecedents	consequents	support	confidence	lift
0	(Cake)	(Coffee)	0.054728	0.526958	1.101515
1	(Cookies)	(Coffee)	0.028209	0.518447	1.083723
2	(Hot chocolate)	(Coffee)	0.029583	0.507246	1.060311
3	(Juice)	(Coffee)	0.020602	0.534247	1.116750
4	(Medialuna)	(Coffee)	0.035182	0.569231	1.189878
5	(Pastry)	(Coffee)	0.047544	0.552147	1.154168
6	(Sandwich)	(Coffee)	0.038246	0.532353	1.112792
7	(Toast)	(Coffee)	0.023666	0.704403	1.472431

Gambar 3.6 Hasil Analisis *Apriori*

Dari hasil implementasi algoritma *Apriori* dengan minimum *support* = 0.02 dan minimum *confidence* = 0.5, ditemukan 8 aturan asosiasi. Proses ini berjalan dengan waktu eksekusi 0.0227 detik dan menggunakan memori 0.74 MB. Dari tabel ini, dapat dilihat bahwa produk *Coffee* adalah produk dominan yang sering muncul sebagai konsekuen dalam aturan asosiasi, yang menunjukkan bahwa pembelian *Coffee* sangat berkaitan dengan pembelian

produk lainnya seperti *Cake*, *Cookies*, dan *Pastry*. Aturan dengan *lift* tertinggi ditemukan pada pola (*Toast* $\rightarrow$ *Coffee*), yang memiliki nilai *lift* 1.4724, mengindikasikan bahwa pembelian *Toast* hampir 47% lebih meningkatkan kemungkinan pembelian *Coffee* dibandingkan dengan ekspektasi acak.

3.3.2 Algoritma *FP-Growth*

Algoritma *FP-Growth* merupakan pendekatan efisien untuk menemukan frequent itemset tanpa perlu menghasilkan kandidat itemset eksplisit seperti pada *Apriori*. Algoritma ini bekerja dengan membangun struktur data pohon yang disebut *FP-Tree* (*Frequent Pattern Tree*), yang mengompresi informasi transaksi dan memungkinkan ekstraksi pola secara rekursif menggunakan pendekatan *divide and conquer*. Tahapan utama dalam algoritma *FP-Growth* meliputi:

1. Membangun *FP-Tree* dari dataset berdasarkan prioritas *item*.
2. Menentukan *conditional pattern base* untuk setiap *item frequent*.
3. Membangun *conditional FP-Tree*.
4. Menghasilkan *frequent itemset* dari pola yang ditemukan.

FP-Growth tidak secara eksplisit menghitung nilai *confidence* dan *lift* pada saat awal, tetapi *itemset* yang dihasilkan kemudian dapat dievaluasi menggunakan rumus yang sama seperti pada algoritma *Apriori*.

Kelebihan *FP-Growth* adalah mampu mengurangi jumlah iterasi dan *overhead* dari pembangkitan kandidat *itemset* seperti pada *Apriori*. Namun, *FP-Growth* cenderung membutuhkan memori lebih besar untuk menyimpan struktur pohon, terutama pada *dataset* yang sangat kompleks atau memiliki banyak item unik.

a. Ilustrasi Perhitungan Manual Algoritma *FP-Growth* dengan Data Sampel

Digunakan dataset sampel yang sama seperti pada *Apriori* (10 transaksi), dan hanya item yang memenuhi minimum *support* (≥ 2) yang dipertahankan. Prioritas item ditentukan berdasarkan frekuensi. Urutan Item Berdasarkan Prioritas:

Transaksi disusun ulang berdasarkan urutan prioritas, dan *FP-Tree* dibangun langkah demi langkah. *Conditional pattern base* dan *conditional FP-Tree* dibentuk untuk masing-masing *item frequent*. Berikut adalah contoh hasilnya:

Mencari nilai prioritas:

Tabel 3.8 Nilai item Prioritas *FP-Growth*

Item	Frekuensi	Prioritas
Bread	4	1
Scandinavian	2	5
Muffin	3	3
Coffee	2	6
Pastry	4	2
Medialuna	3	4

Mengurutkan berdasarkan prioritas:

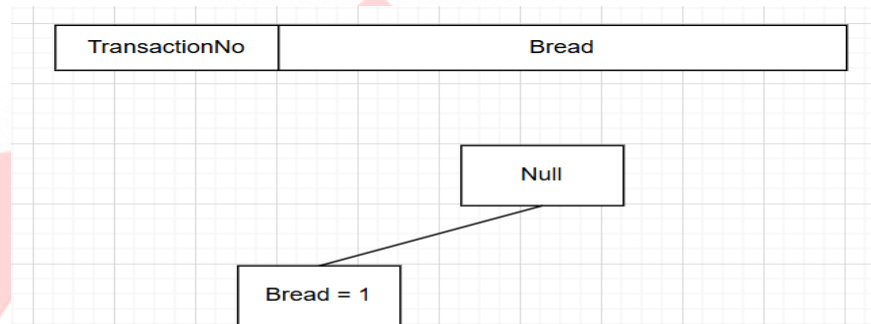
Tabel 3.9 Mengurutkan item prioritas *FP-Growth*

Transaction ID	Items
1	Bread
2	Scandinavian
3	- (tidak memenuhi syarat)
4	Muffin
5	Bread, Pastry, Coffee
6	Pastry, Muffin, Medialuna
7	Pastry, Medialuna, Coffee
8	Bread, Pastry
9	Bread, Muffin
10	Medialuna, Scandinavian

Selanjutnya, membuat *FP-Tree* dari data sebelumnya:

1. Langkah 1 – Transaksi ID 1: *Bread*

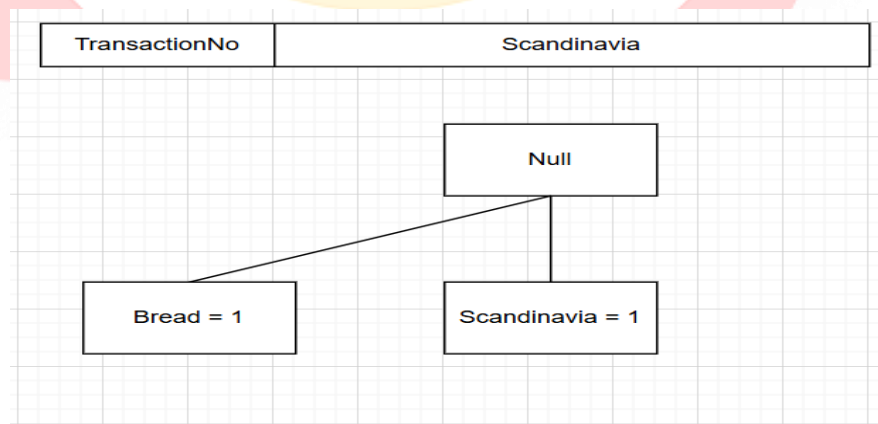
Pada langkah ini, item "*Bread*" dari transaksi pertama dimasukkan ke dalam *FP-Tree* sebagai *child* dari *node root (Null)* dengan *count* = 1.



Gambar 3.7 Langkah Pertama Pembentukan *FP-Tree*

2. Langkah 2 – Transaksi ID 2: *Scandinavian*

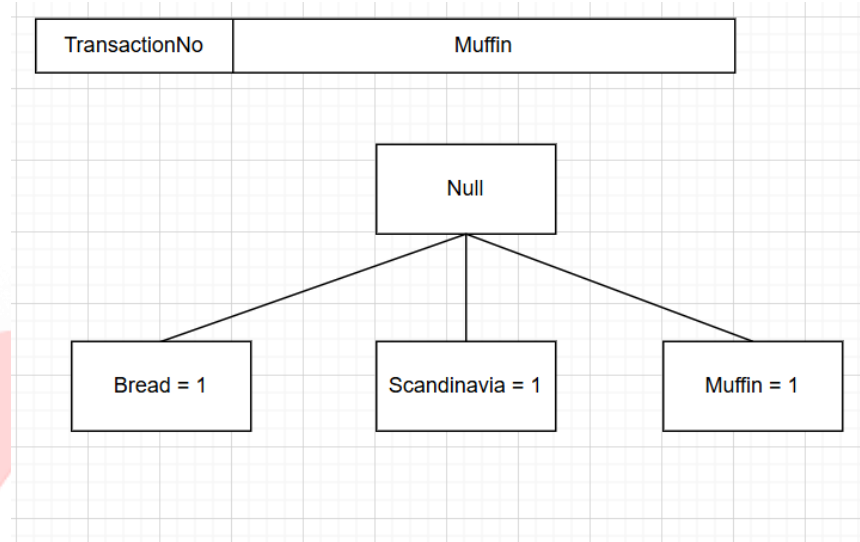
Item "*Scandinavian*" dimasukkan sebagai *child* baru dari *root* karena belum ada sebelumnya, dengan *count* = 1.



Gambar 3.8 Langkah Kedua Pembentukan *FP-Tree*

3. Langkah 4 – Transaksi ID 4: *Muffin*

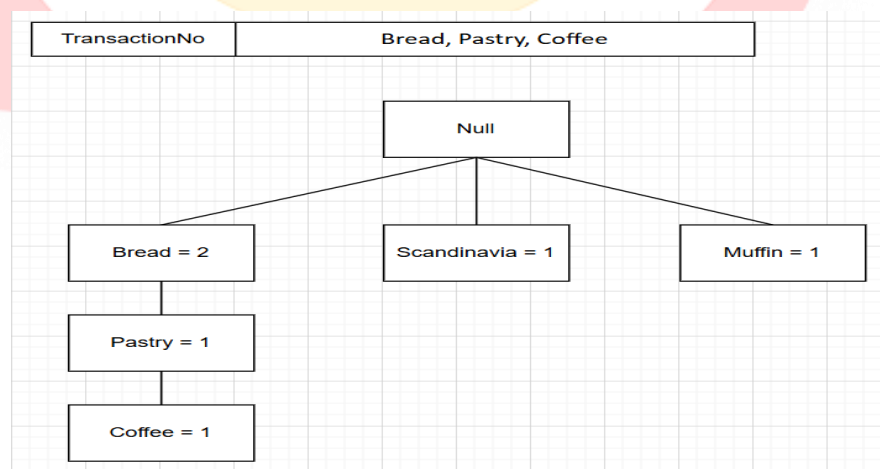
Item "*Muffin*" dimasukkan sebagai *child* baru dari *root*, dengan *count* = 1.



Gambar 3.9 Langkah Ketiga Pembentukan *FP-Tree*

4. Langkah 5 – Transaksi ID 5: *Bread, Pastry, Coffee*

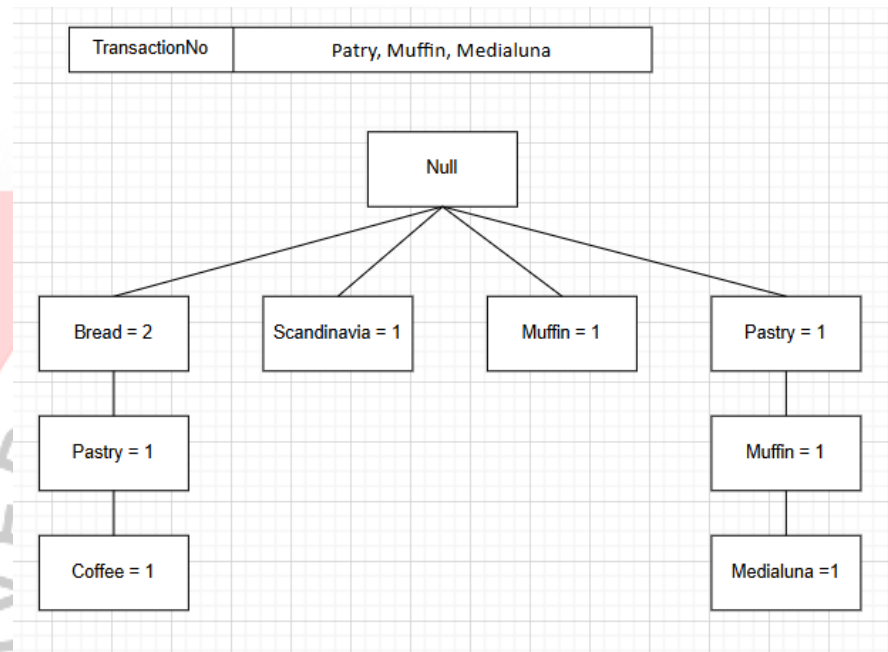
Item "*Bread*" sudah ada, maka *count*-nya ditambah menjadi 2. "*Pastry*" dimasukkan sebagai *child* dari "*Bread*". "*Coffee*" dimasukkan sebagai *child* dari "*Pastry*".



Gambar 3.10 Langkah Keempat Pembentukan *FP-Tree*

5. Langkah 6 – Transaksi ID 6: *Pastry, Muffin, Medialuna*

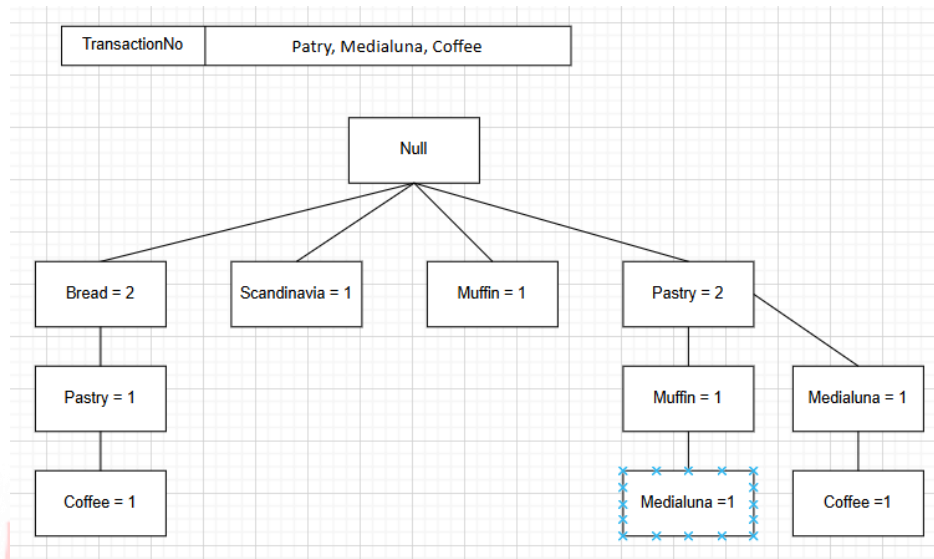
"*Pastry*" dimasukkan sebagai *child* baru dari *root*. "*Muffin*" dimasukkan sebagai *child* dari "*Pastry*". "*Medialuna*" dimasukkan sebagai *child* dari "*Muffin*".



Gambar 3.11 Langkah Kelima Pembentukan *FP-Tree*

6. Langkah 7 – Transaksi ID 7: *Pastry, Medialuna, Coffee*

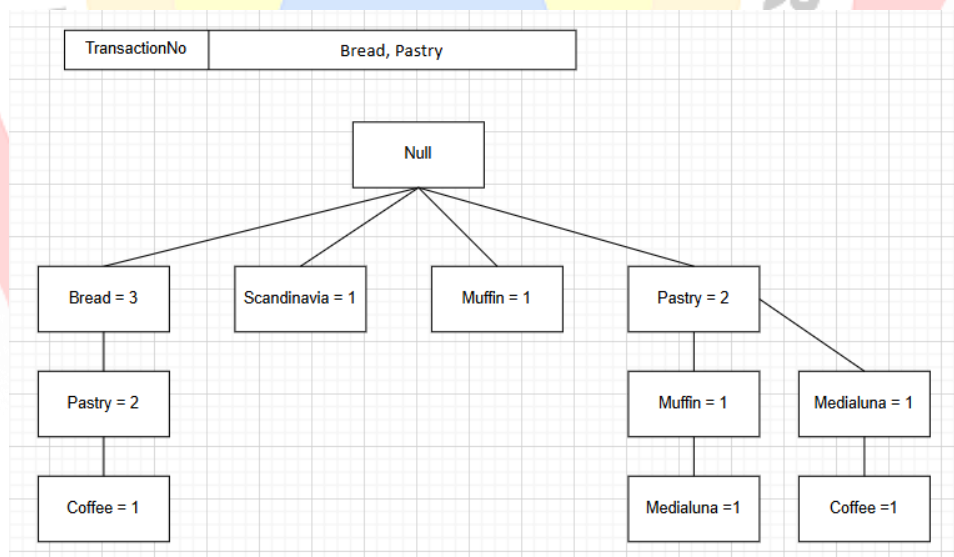
"*Pastry*" sudah ada sebagai *child* *root*, *count* ditambah jadi 2. "*Medialuna*" ditambahkan sebagai *child* dari "*Pastry*". "*Coffee*" ditambahkan sebagai *child* dari "*Medialuna*".



Gambar 3.12 Langkah Keenam Pembentukan *FP-Tree*

7. Langkah 8 – Transaksi ID 8: *Bread*, *Pastry*

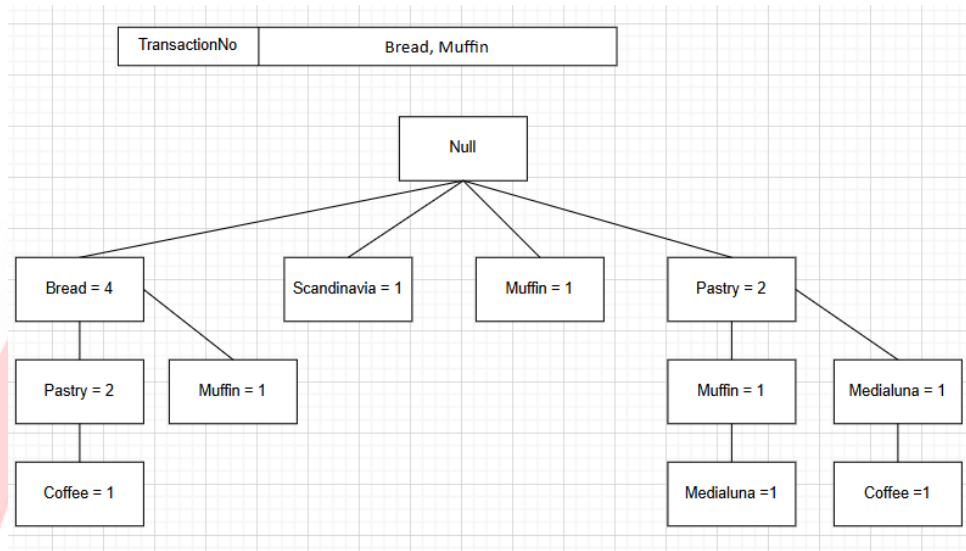
"*Bread*" sudah ada, *count* ditambah jadi 3. "*Pastry*" sudah ada sebagai *child* dari "*Bread*", *count* ditambah jadi 2.



Gambar 3.13 Langkah Ketujuh Pembentukan *FP-Tree*

8. Langkah 9 – Transaksi ID 9: *Bread*, *Muffin*

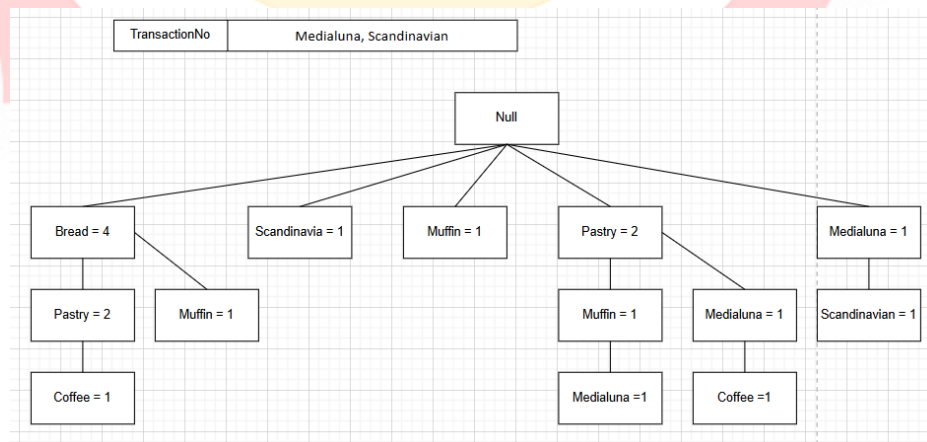
"*Bread*" sudah ada, *count* ditambah jadi 4. "*Muffin*" dimasukkan sebagai *child* dari "*Bread*".



Gambar 3.14 Langkah Kedelapan Pembentukan *FP-Tree*

9. Langkah 10 – Transaksi ID 10: *Medialuna*, *Scandinavian*

"*Medialuna*" dimasukkan sebagai *child* baru dari *root*. "*Scandinavian*" dimasukkan sebagai *child* dari "*Medialuna*".



Gambar 3.15 Langkah Kesembilan Pembentukan *FP-Tree*

Transaksi disusun ulang berdasarkan urutan prioritas, dan *FP-Tree* dibangun langkah demi langkah. *Conditional pattern base* dan *conditional FP-Tree* dibentuk untuk masing-masing *item frequent*. Berikut adalah hasilnya:

Frequent Pattern Generate:

Tabel 3.10 Frequent Pattern Generate

Items	Support
Coffee, Pastry	2
Medialuna, Pastry	2
Pastry, Bread	2

Aturan asosiasi yg lolos:

Tabel 3.11 Aturan Asosiasi FP-Growth

Rule	Confidence(%)	Lolos Min. Confidence
Coffee , Pastry	100%	Ya
Medialuna, Pastry	66,7%	Ya
Pastry, Bread	50%	Tidak

- b. Implementasi dan Analisis Algoritma *FP-Growth* dengan Dataset bakery.csv (18.888 Transaksi)

Pada eksperimen yang sama, algoritma *FP-Growth* juga menghasilkan 8 aturan asosiasi dengan waktu eksekusi 0.4035 detik dan memori 2.07 MB. Walaupun memerlukan waktu yang lebih lama dan memori yang lebih besar dibandingkan dengan *Apriori*, *FP-Growth* tetap menghasilkan hasil yang sangat mirip dengan *Apriori*. Berikut adalah semua aturan yang ditemukan menggunakan *FP-Growth*:

=== FP-GROWTH RESULTS ===

Jumlah aturan : 8
Waktu eksekusi : 0.4035 detik
Penggunaan memori : 2.07 MB

Semua aturan yang ditemukan:

	antecedents	consequents	support	confidence	lift
0	(Hot chocolate)	(Coffee)	0.029583	0.507246	1.060311
1	(Cookies)	(Coffee)	0.028209	0.518447	1.083723
2	(Pastry)	(Coffee)	0.047544	0.552147	1.154168
3	(Medialuna)	(Coffee)	0.035182	0.569231	1.189878
4	(Juice)	(Coffee)	0.020602	0.534247	1.116750
5	(Cake)	(Coffee)	0.054728	0.526958	1.101515
6	(Sandwich)	(Coffee)	0.038246	0.532353	1.112792
7	(Toast)	(Coffee)	0.023666	0.704403	1.472431

Gambar 3.16 Hasil Analisis *FP-Growth*

Dari hasil ini, dapat dilihat bahwa aturan yang ditemukan oleh *FP-Growth* sangat mirip dengan *Apriori*, dengan *Coffee* sebagai konsekuen dominan dalam banyak aturan. *Lift* tertinggi pada (*Toast* → *Coffee*) menunjukkan bahwa hubungan antara *Toast* dan *Coffee* juga sangat kuat pada algoritma *FP-Growth*, sama seperti pada *Apriori*. Hal ini mengindikasikan konsistensi hasil yang baik antara kedua algoritma.

3.3.3 Analisis Perbandingan Kinerja Algoritma

Berdasarkan hasil implementasi dan analisis terhadap algoritma *Apriori* dan *FP-Growth* menggunakan *dataset* Bakery.csv yang terdiri atas 9.465 transaksi unik, diperoleh perbandingan kinerja kedua algoritma seperti ditunjukkan pada tabel berikut:

Tabel 3.12 Hasil Perbandingan *Apriori* dan *FP-Growth*

Aspek	<i>Apriori</i>	<i>FP-Growth</i>
Jumlah aturan	8	8
Waktu eksekusi	0.0227 detik	0.4035 detik
Penggunaan memori	0.74 MB	2.07 MB
Lift tertinggi	1.1899	1.1899
Pola dominan	Coffee	Coffee
Aturan paling kuat	Medialuna → Coffee	Medialuna → Coffee

1) Waktu Eksekusi

Algoritma *Apriori* menunjukkan keunggulan dari segi waktu eksekusi, dengan waktu hanya 0.0227 detik untuk mengekstraksi aturan asosiasi. Ini menjadikannya pilihan yang sangat efisien untuk *dataset* dengan ukuran yang lebih kecil dan kurang *kompleks*. Di sisi lain, *FP-Growth* memerlukan waktu eksekusi lebih lama, yaitu 0.4035 detik. Hal ini karna mengingat *FP-Growth* menggunakan struktur *FP-Tree* yang lebih *kompleks*, yang memungkinkan algoritma ini menangani *dataset* yang lebih besar dengan lebih efisien meskipun memerlukan waktu yang lebih lama pada *dataset* yang lebih kecil.

2) Penggunaan Memori

Dari segi penggunaan memori, *Apriori* membutuhkan memori lebih sedikit, yaitu 0.74 MB, menjadikannya pilihan yang lebih hemat memori untuk aplikasi dengan sumber daya terbatas. Sebaliknya, *FP-Growth* menggunakan memori lebih besar, yaitu 2.07 MB. Hal ini disebabkan oleh pemrosesan struktur *FP-Tree*, yang menyimpan lebih banyak informasi selama proses mining aturan asosiasi.

3) Kualitas Aturan

Kedua algoritma menghasilkan aturan dengan *lift* tertinggi pada aturan *Medialuna* → *Coffee*, yang memiliki nilai 1.1899 pada keduanya. Nilai *lift* ini menunjukkan bahwa pembelian *Medialuna* meningkatkan kemungkinan pembelian *Coffee* sekitar 19% lebih tinggi dibandingkan dengan ekspektasi acak, yang menandakan adanya hubungan yang kuat antara kedua produk tersebut. Selain itu, *Coffee* muncul sebagai item yang paling sering dibeli bersama produk

lain di toko *bakery*, baik pada algoritma *Apriori* maupun *FP-Growth*, yang menunjukkan bahwa *Coffee* adalah produk yang sangat terkait dengan produk lainnya di toko tersebut.

4) Rekomendasi Algoritma

Berdasarkan hasil eksperimen dengan dataset *bakery.csv*, dapat disimpulkan bahwa *Apriori* lebih cocok digunakan untuk analisis *dataset* kecil hingga menengah, seperti yang digunakan dalam penelitian ini. Dengan waktu eksekusi yang sangat cepat dan penggunaan memori yang efisien, *Apriori* adalah pilihan yang tepat untuk analisis yang memerlukan kecepatan dan hemat sumber daya. Oleh karena itu, *Apriori* sangat cocok digunakan untuk analisis pola pembelian konsumen dalam skala kecil atau aplikasi yang membutuhkan hasil cepat dengan batasan memori dan waktu.

3.4 Kerangka Pemikiran

Kerangka pemikiran ini menggambarkan alur yang digunakan untuk mencapai solusi dari masalah yang ada dalam penelitian ini, yaitu penerapan algoritma *Apriori* dan *FP-Growth* untuk *Market Basket Analysis* pada data transaksi penjualan *bakery*.

KONDISI SAAT INI :

- Data penjualan bakery tersedia dalam bentuk transaksi harian.
- Belum dilakukan analisis mendalam terkait pola pembelian (market basket analysis).
- Algoritma Apriori dan FP-Growth sering digunakan untuk Market Basket Analysis tetapi memiliki kelebihan dan kekurangan masing-masing.

INDIKATOR :

- Waktu Komputasi Algoritma: Mengukur kecepatan eksekusi algoritma dalam menyelesaikan analisis pola pembelian pada dataset.
- Penggunaan memori: Seberapa efisien algoritma memanfaatkan memori selama proses analisis.
- Akurasi dalam Menghasilkan Aturan Asosiatif: Menilai kemampuan algoritma dalam menghasilkan pola atau hubungan antarproduk yang relevan dan dapat diandalkan.

SISTEM YG DIUSULKAN :



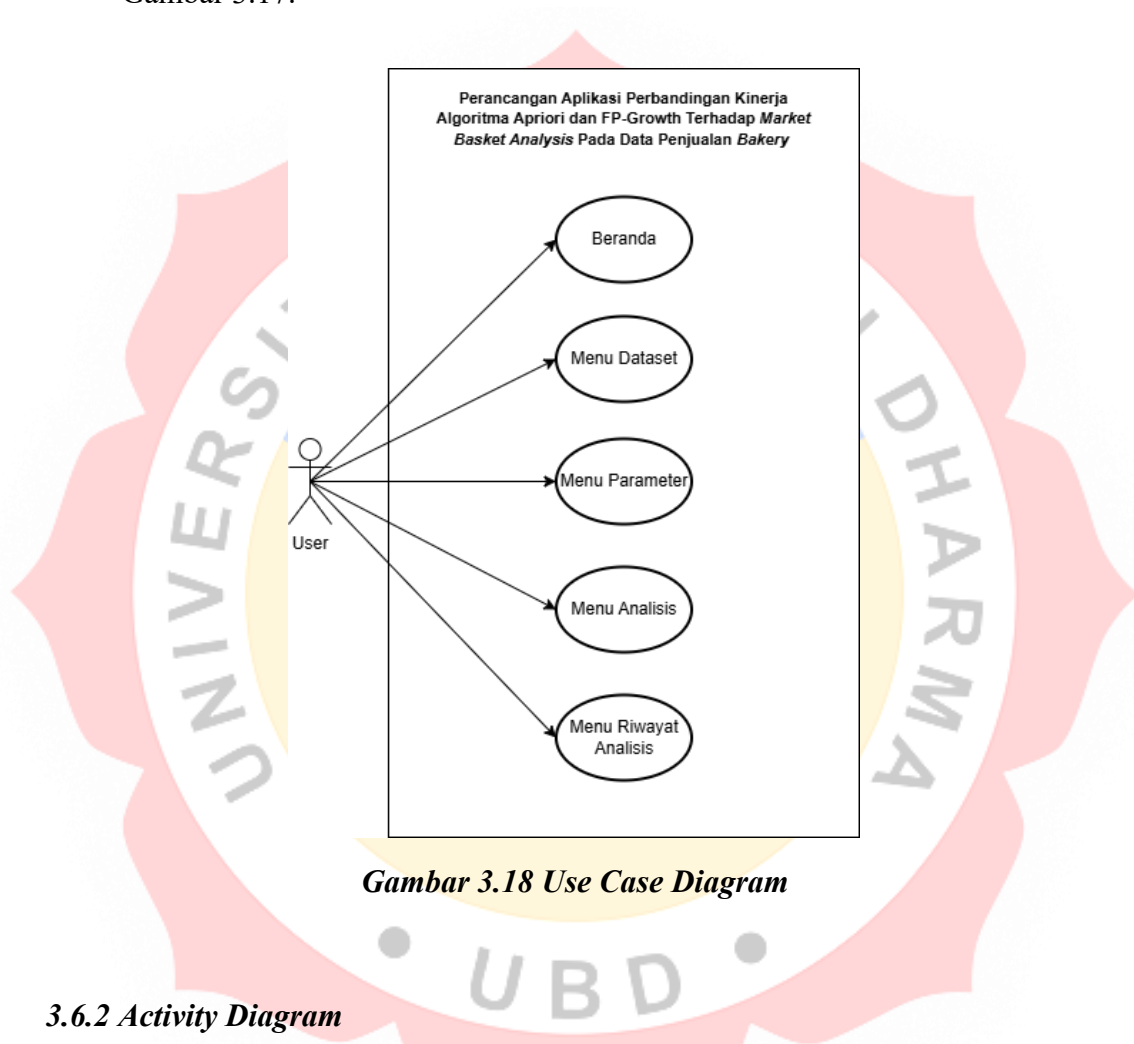
Gambar 3.17 Kerangka Pemikiran

3.5 Perancangan UML

Pada fase ini, akan dilakukan perancangan terhadap sistem yang telah diusulkan. Perancangan sistem ini dapat direpresentasikan menggunakan *Unified Modeling Language (UML)*, yang mencakup beberapa diagram seperti *Use Case Diagram*, *Class Diagram*, *Activity Diagram*, dan *Sequence Diagram*.

3.6.1 Use Case Diagram

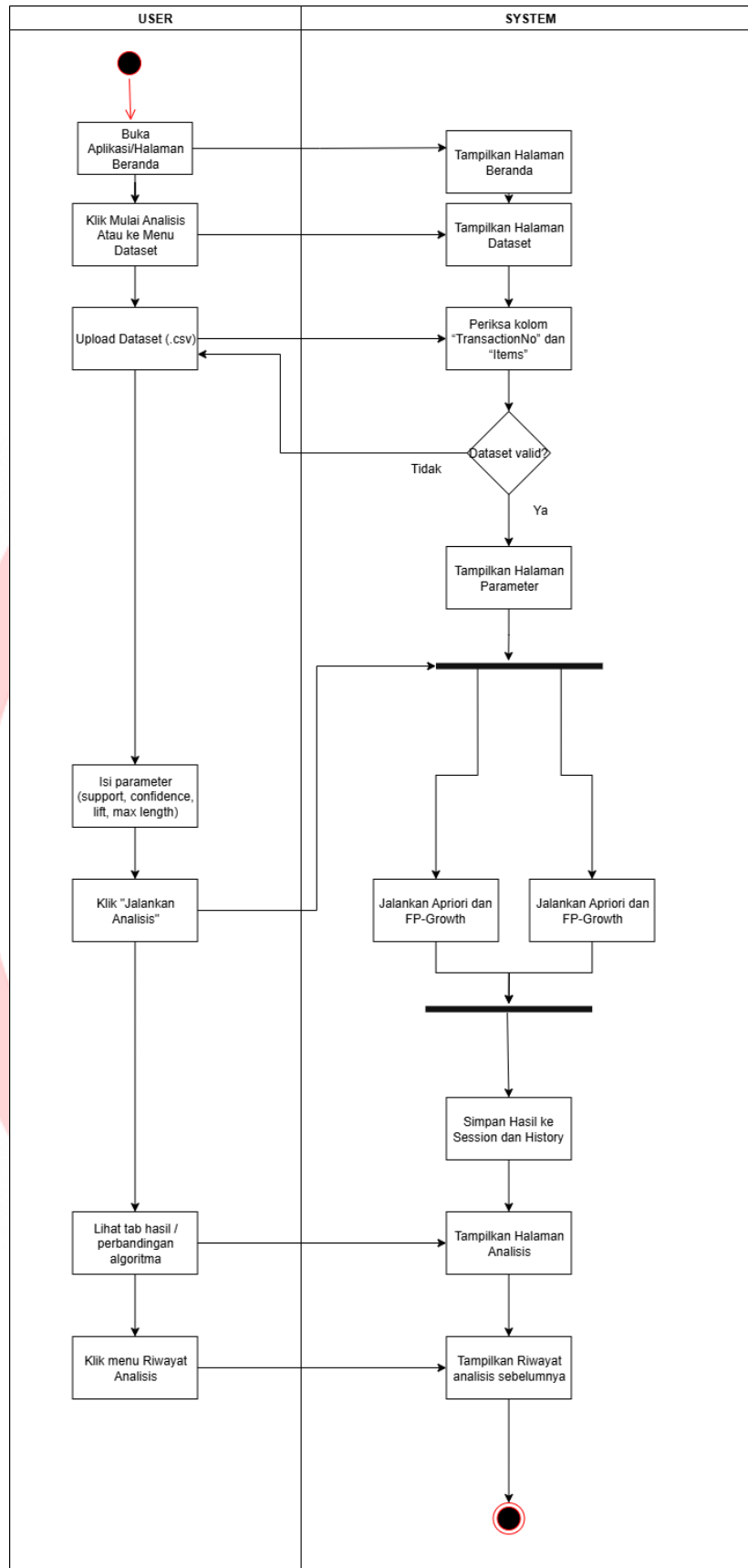
Dalam pengembangan suatu program, dibutuhkan sebuah model data berupa diagram yang dapat menggambarkan alur proses sistem yang akan dibangun. Berikut adalah bentuk diagram *Use Case* yang dapat dilihat pada Gambar 3.17.



Gambar 3.18 Use Case Diagram

3.6.2 Activity Diagram

Rangkuman kegiatan pada setiap *event* sistem digambarkan pada *activity diagram* dapat dilihat pada gambar 3.16 berikut:



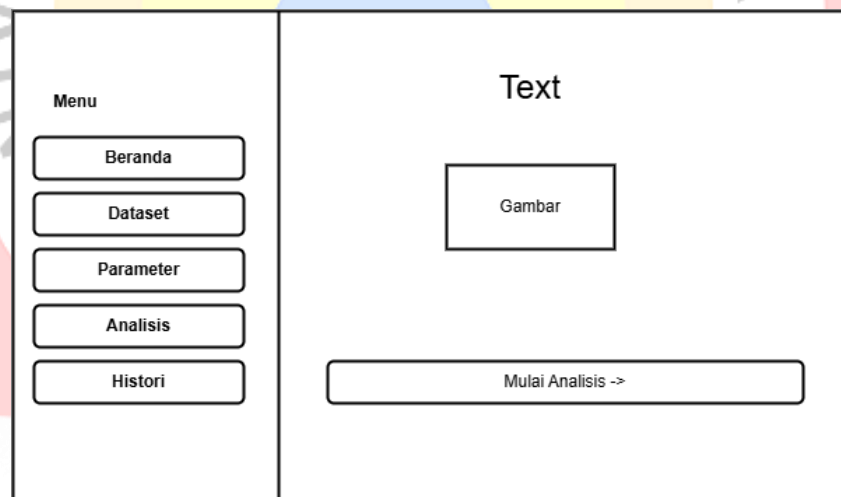
Gambar 3.19 Activity Diagram

3.6 Perancangan Layar dan Menu

Perancangan layar, menu, dan database bertujuan untuk mendukung proses analisis perbandingan algoritma *Apriori* dan *FP-Growth* secara terstruktur dan efisien. Sistem ini dibangun dengan beberapa halaman utama, yaitu Halaman Beranda, Halaman Dataset, Halaman *Parameter*, Halaman Analisis, dan Halaman *Histori*. Berikut adalah uraian perancangan untuk masing-masing bagian:

1. Rancangan Halaman Beranda

Halaman ini merupakan tampilan awal dari sistem yang berisi informasi pengantar dan *navigasi* ke menu utama. Desainnya bersifat ringkas dan langsung mengarahkan ke proses inti sistem. Halaman ini juga berfungsi sebagai titik awal sebelum melanjutkan ke proses unggah dataset dan analisis, seperti pada gambar 3.25 berikut ini.

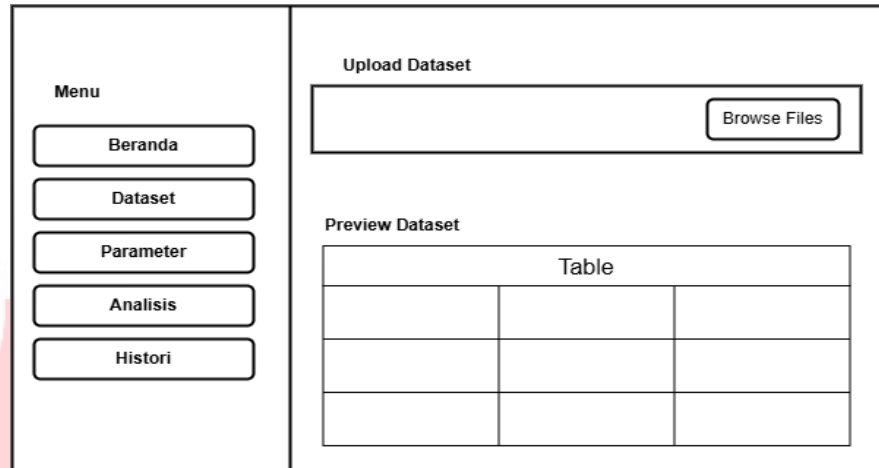


Gambar 3.20 Rancangan Halaman Beranda

2. Rancangan Menu *Dataset*

Halaman pada menu ini menyediakan fitur untuk mengunggah dataset yang akan digunakan dalam proses analisis. Setelah proses unggah berhasil, sistem akan memunculkan tampilan visual dari isi dataset, seperti tabel data yang telah

dimuat. Tujuan dari tampilan ini adalah untuk memastikan bahwa data telah dimasukkan dengan benar sebelum dianalisis lebih lanjut dapat dilihat pada gambar 3.27



Menu

Beranda

Dataset

Parameter

Analisis

Histori

Upload Dataset

Browse Files

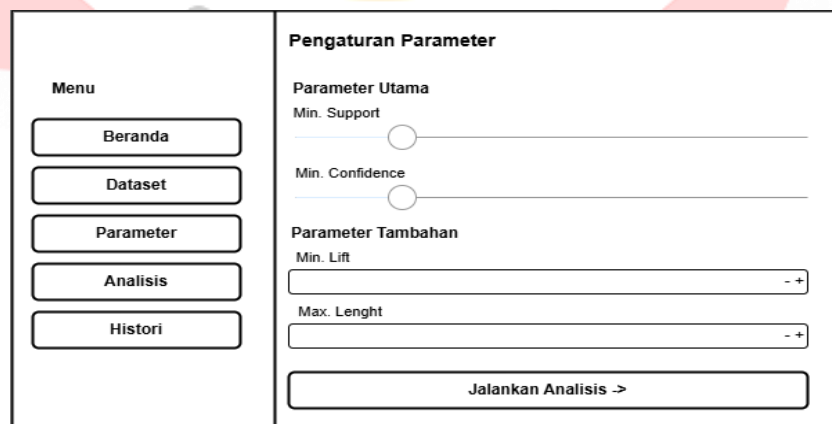
Preview Dataset

Table

Gambar 3.21 Rancangan Menu *Dataset*

3. Rancangan Menu Parameter

Halaman menu parameter memungkinkan pengaturan nilai-nilai penting sebelum proses analisis dimulai, seperti minimum *support* dan minimum *confidence*. Pengaturan ini akan memengaruhi hasil dari algoritma *Apriori* dan *FP-Growth*. Rancangan halaman ini menampilkan kolom input yang dapat diisi secara manual sesuai kebutuhan analisis dapat dilihat pada gambar 3.28



Menu

Beranda

Dataset

Parameter

Analisis

Histori

Pengaturan Parameter

Parameter Utama

Min. Support

Min. Confidence

Parameter Tambahan

Min. Lift

Max. Length

Jalankan Analisis ->

Gambar 3.22 Rancangan Menu Parameter

4. Rancangan Menu Analisis

Halaman menu analisis dibagi menjadi tiga bagian utama. Bagian pertama adalah bagian perbandingan, yang menampilkan perbedaan kinerja antara algoritma *Apriori* dan *FP-Growth*, baik dari sisi waktu eksekusi maupun jumlah aturan yang dihasilkan. Bagian kedua adalah bagian *apriori*, yang menunjukkan hasil aturan asosiasi berdasarkan algoritma *Apriori*. Sementara itu, bagian ketiga adalah bagian *fp-growth*, yang menampilkan hasil aturan asosiasi dari algoritma *FP-Growth*. Setiap bagian disajikan dalam format tabel yang jelas dan mudah dianalisis dapat dilihat pada gambar 3.29

Menu

- Beranda
- Dataset
- Parameter
- Analisis
- Histori

Analisis Algoritma

Perbandingan Apriori FP-Growth

Perbandingan Kinerja

Rekomendasi Algoritma :
Text

Grafik Perbandingan

Lanjut ke Histori ->

Gambar 3.23 Rancangan Menu Analisis

5. Rancangan Menu *Histori*

Halaman *histori* menyimpan rekam jejak hasil analisis yang telah dilakukan sebelumnya. Halaman ini terdiri dari tiga bagian: bagian ringkasan, yang menampilkan daftar *histori* analisis; bagian *parameter*, yang memperlihatkan nilai *parameter* yang digunakan dalam setiap analisis; dan bagian *detail*, yang menyajikan informasi lengkap dari hasil aturan asosiasi yang dihasilkan. Dengan adanya halaman ini, hasil analisis terdahulu dapat ditinjau kembali tanpa perlu mengulangi proses analisis dari awal dapat dilihat pada gambar 3.30

Menu

Beranda

Dataset

Parameter

Analisis

Histori

Histori Analisis
Analisis Sebelumnya

Ringkasan

Parameter

Detail

Hapus Semua Histori

Kembali ke Beranda ->

Gambar 3.24 Rancangan Menu *Histori*

3.7 Jadwal Penelitian

Jadwal penelitian disusun untuk mengatur alur kerja secara terstruktur agar setiap tahap penelitian dapat diselesaikan tepat waktu. Berikut adalah tahapan dan jadwal pelaksanaan penelitian, yang terbagi menjadi beberapa aktivitas utama:

Tabel 3.13 Jadwal Penelitian

No	Kegiatan	Bulan															
		MARET 2025				APRIL 2025				MEI 2025				JUNI 2025			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	BAB 1																
2	REVISI PROGRAM																
3	STUDI LITERATUR																
4	BAB 2																
5	BAB 3																
6	BAB 4																
7	TESTING PROGRAM																
8	BAB 5																