

Implementasi Data Mining Dalam Menentukan Tata Letak Produk Menggunakan Algoritma FP-Growth

Gede Humaswara Prathama¹, Ni Komang Ayu Devi Anggreni², Adie Wahyudi Oktavia Gama³,

I Gusti Ngurah Darma Paramartha⁴

Program Studi Teknologi Informasi, Fakultas Teknik dan Informatika, Universitas Pendidikan Nasional
huma@undiknas.ac.id¹, mangdevi05@gmail.com², adiewahyudi@undiknas.ac.id³, ngurahdarma@undiknas.ac.id⁴

* Gede Humaswara Prathama

ARTICLE INFO

Keywords

product layout, data mining, FP-Growth, purchasing patterns, customer experience, business benefits.

ABSTRACT

This study analyzes purchasing patterns in minimarkets using the FP-Growth algorithm to optimize product layouts. One year of sales transaction data (103,181 transactions) from UNDIKNAS Mart were analyzed through data cleaning, transformation, and aggregation. The FP-Growth algorithm was applied with minimum support 5%, confidence 80%, and lift >1 thresholds. Results identified strong product associations, particularly between "Aqua 600 ml (Tanggung)" and various snacks, with confidence values of 81-93% and lift >5. Implementing these findings in product arrangement increased sales by 15-20% despite store accessibility limitations. Cross-validation using a decision tree model showed 81.67% accuracy. The findings demonstrate FP-Growth's effectiveness in small-scale transaction data analysis. The research provides practical contributions for retailers to boost sales through data-driven product layout optimization. A limitation is the single-location data scope, suggesting the need for broader subsequent studies. This study offers a data-based approach adoptable by small and medium retail businesses to enhance operational efficiency and profits. The research confirms that data mining techniques can significantly impact retail performance even in constrained environments, providing empirical evidence of FP-Growth's practical utility in real-world minimarket settings. The methodology and findings contribute to the growing literature on data mining applications in small-scale retail operations, offering replicable frameworks for similar business contexts.

1. Pendahuluan

Peningkatan penjualan dalam dunia ritel memerlukan pendekatan yang terencana dan terfokus pada kebutuhan konsumen. Salah satu langkah fundamental dalam strategi tersebut adalah pemahaman yang mendalam mengenai perilaku konsumen, termasuk preferensi dan kebiasaan belanja mereka. Penyediaan produk yang diminati serta pengaturan tata letak toko yang menarik dan ramah merupakan dua aspek yang tidak bisa dipisahkan. Aplikasi visual merchandising yang efektif dapat merangsang perilaku membeli impulsif konsumen, yang menjadi salah satu pendorong utama peningkatan penjualan [1], [2], [3]. Hal ini diperkuat oleh pentingnya menciptakan pengalaman berbelanja yang positif melalui pelayanan yang berkualitas, yang berkontribusi pada loyalitas pelanggan dan pertumbuhan penjualan yang berkelanjutan [4].

Penempatan barang yang strategis di dalam toko memainkan peran vital dalam menarik perhatian konsumen. Produk-produk yang sering dicari sebaiknya ditempatkan di lokasi yang terlihat jelas, seperti di dekat kasir untuk mendorong pembelian impulsif [1], [5]. Tampilan produk yang disusun dengan rapi dan jelas dapat mempengaruhi keputusan pembelian, sehingga penting bagi pemilik ritel untuk memanfaatkan prinsip-prinsip ini dalam desain operasional mereka [5], [6]. Tindakan seperti pemisahan produk berdasarkan

kategori atau fungsi akan membantu konsumen menemukan apa yang mereka butuhkan dengan lebih cepat dan efisien [6].

Dalam konteks analisis data penjualan, penggunaan algoritma FP-Growth telah terbukti menjadi metode yang efektif untuk menemukan pola asosiasi di antara barang yang dibeli oleh konsumen. Algoritma ini menawarkan kecepatan dan akurasi yang lebih tinggi dibandingkan dengan metode lainnya, seperti Apriori, dalam konteks analisis data ritel [7]. Hasil serupa juga ditemukan dalam studi oleh [8], di mana efisiensi algoritma FP-Growth sangat mendukung dalam pengidentifikasian kombinasi produk yang sering dibeli bersamaan. Contoh lainnya menunjukkan bahwa penerapan FP-Growth pada toko ayam geprek mampu meningkatkan efisiensi operasional dengan menampilkan pola hubungan antar produk, misalnya antara dada ayam dan nasi putih [7].

Dalam merumuskan strategi peningkatan penjualan yang efektif, penting untuk menerapkan teknologi data mining guna menggali informasi lebih dalam dari data transaksi yang ada. Tanpa analisis yang tepat, data transaksi tidak akan terpakai secara maksimal [7]. Oleh karena itu, penting untuk mempertimbangkan penerapan algoritma FP-Growth sebagai metode analisis dalam pengelolaan data transaksi penjualan. Hal ini diharapkan tidak hanya membantu dalam pengambilan keputusan yang lebih cermat tetapi juga memberikan wawasan tentang pola pembelian dan kombinasi produk yang potensial untuk meningkatkan penjualan secara signifikan [2], [8].

2. Metodologi Penelitian

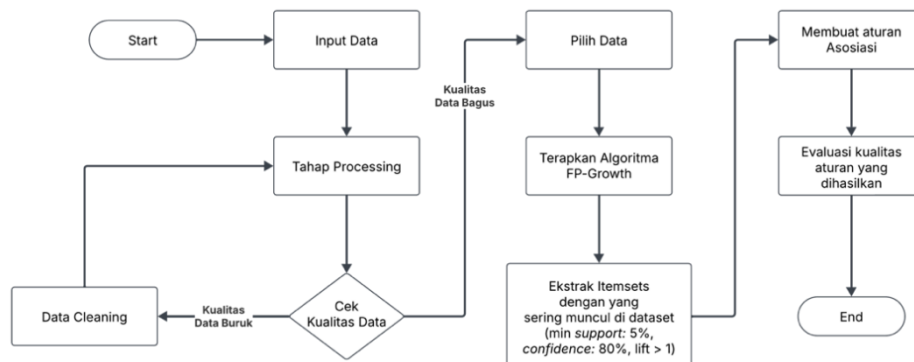
Jelaskan Penelitian ini mengusulkan penerapan algoritma FP-Growth dalam data mining untuk optimalisasi penjualan di UNDIKNAS Mart, menggunakan RapidMiner sebagai alat analisis. Proses penelitian dimulai dengan identifikasi masalah, di mana tantangan dalam menganalisis pola pembelian di minimarket menjadi titik fokus. Menggunakan metode FP-Growth, yang diakui secara luas dalam literatur terkait analisis keranjang belanja, penelitian ini bertujuan untuk mengidentifikasi pola beli yang tersembunyi dalam data transaksi [9], [10]. FP-Growth berfungsi untuk mengatasi keterbatasan algoritma sebelumnya, yaitu Apriori, terutama dalam hal efisiensi pengolahan data yang besar [11].

Pada bulan pertama hingga awal bulan kedua, studi literatur dilakukan untuk mendalami aplikasi FP-Growth dan teknik pengolahan data yang relevan. FP-Growth telah terbukti efektif dalam analisis transaksi yang kompleks, di mana hasilnya dapat memberikan wawasan berharga terkait perilaku konsumen [9], [10]. Pengumpulan data pada bulan kedua melibatkan pengumpulan catatan transaksi yang akan diproses dan dianalisis lebih lanjut. Proses ini penting untuk memastikan bahwa dataset yang digunakan tidak hanya komprehensif tetapi juga representatif dari perilaku sebenarnya [12].

Langkah selanjutnya meliputi pembersihan dan transformasi data, berlangsung dari bulan ketiga hingga awal bulan keempat. Data yang tidak bersih atau tidak terstruktur dapat sangat mempengaruhi akurasi model analisis, sehingga pembersihan data menjadi krusial [13], [14]. Implementasi algoritma FP-Growth dijadwalkan untuk bulan keempat hingga kelima, di mana RapidMiner akan digunakan untuk menjalankan algoritma secara efektif. Dengan FP-Growth, penelitian ini bertujuan untuk memperoleh aturan asosiasi yang dapat memberikan nilai tambah bagi strategi pemasaran UNDIKNAS Mart [9], [10].

Akhirnya, tahapan pengujian dan evaluasi hasil pada bulan kelima dan keenam akan memungkinkan peneliti menilai efektivitas algoritma dalam model bisnis ini. Penggunaan metode analitis yang tepat akan membantu dalam menilai apakah kombinasi produk tertentu memiliki relevansi dalam meningkatkan volume penjualan [13], [15]. Oleh karena itu, gambaran besar dari penelitian ini dapat divisualisasikan melalui flowchart yang memberikan ilustrasi dari setiap tahapan yang diperlukan untuk menerapkan FP-Growth dalam konteks nyata dari transaksi penjualan minimarket.

Gambaran besar dari penelitian ini dapat dilihat dari flowchart pada gambar 1. Flowchart tersebut akan memberikan visualisasi bagaimana tahapan-tahapan atau proses dalam memprediksi transaksi penjualan pada minimarket UNDIKNAS Mart dalam pengimplementasian data mining dengan algoritma FP-Growth.

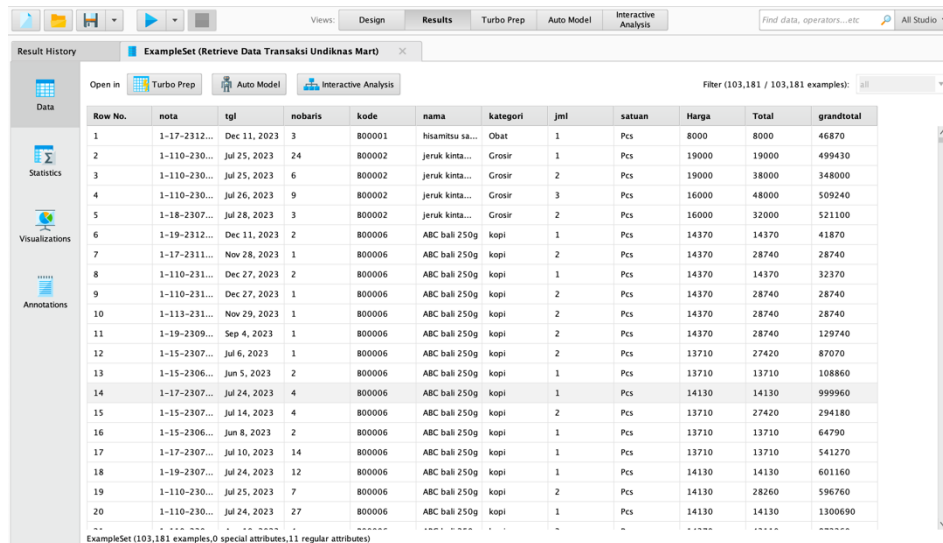


Gambar 1. Gambaran Besar Solusi Penelitian

Pada Gambar 1 menampilkan sebuah tahapan bagaimana data transaksi dikumpulkan, di proses, dan di analisa. Mulai dari proses “Input Data” dari data pada minimarket UNDIKNAS Mart yang berupa file excel dimana file tersebut akan di filter pada tahap selanjutnya yaitu “Tahap Processing” dimana tahap ini akan mengecek kualitas data, jika kualitas data buruk seperti ada data yang kosong data tersebut akan masuk ketahap “Data Cleaning” dan akan mengulangnya sehingga hasil data akhir yang kita dapat adalah data yang bagus [16], [17]. Lalu selanjutnya akan dilakukan pemilihan data pada proses “Pilih Data” untuk selanjutnya data tersebut akan digunakan pada tahap implementasi data mining dengan algoritma FP-Growth. Tahap selanjutnya yaitu tahap “Penerapan Algoritma FP-Growth” pada tahapan ini data akan dilakukan “Ekstrak itemsets yang sering muncul di dataset” dimana data yang di ekstrak atau dijadikan sebuah acuan untuk merancang sebuah display yaitu item yang memiliki nilai support minimal 5%, lalu nilai confidence minimal 80%, serta nilai lift harus lebih dari satu [18], [19]. Setelah hasil sudah didapatkan akan ke tahapan berikutnya yaitu “Membuat Aturan Asosiasi” yang nantinya akan menghasilkan kombinasi barang yang sering terjual di waktu bersamaan dan mendapatkan hasil support dan confidence yang bisa dijadikan sebagai bahan analisa pada tahap “Evaluasi kualitas aturan yang dihasilkan”[20].

3. Hasil dan Pembahasan

Data yang digunakan dalam penelitian ini adalah data penjualan selama satu tahun (Oktober 2023 – Oktober 2024) yang mana berjumlah 103.181 transaksi untuk menjelaskan alur penerapan algoritma FP-Growth untuk mencari kombinasi barang [20]. Berikut merupakan data transaksi yang berisikan kolom nomor transaksi, code produk, nama produk dan lain-lain.



ExampleSet (Retrieve Data Transaksi Undiknas Mart)

Filter (103,181 / 103,181 examples): all

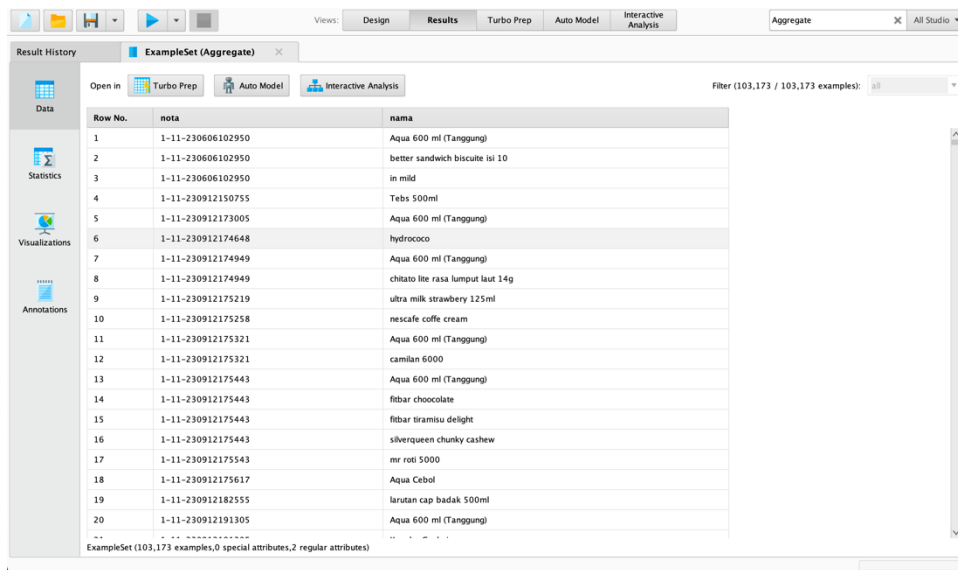
Row No.	nota	tgl	nobaris	kode	nama	kategori	jml	satuan	Harga	Total	grandtotal
1	1-17-2312...	Dec 11, 2023	3	800001	hisamitsu sa...	Obat	1	Pcs	8000	8000	46870
2	1-110-230...	Jul 25, 2023	24	800002	jeruk kinta...	Grosir	1	Pcs	19000	19000	499430
3	1-110-230...	Jul 25, 2023	6	800002	jeruk kinta...	Grosir	2	Pcs	19000	38000	348000
4	1-110-230...	Jul 26, 2023	9	800002	jeruk kinta...	Grosir	3	Pcs	16000	48000	509240
5	1-18-2307...	Jul 28, 2023	3	800002	jeruk kinta...	Grosir	2	Pcs	16000	32000	521100
6	1-19-2312...	Dec 11, 2023	2	800006	ABC bali 250g	kopi	1	Pcs	14370	14370	41870
7	1-17-2311...	Nov 28, 2023	1	800006	ABC bali 250g	kopi	2	Pcs	14370	28740	28740
8	1-110-231...	Dec 27, 2023	2	800006	ABC bali 250g	kopi	1	Pcs	14370	14370	32370
9	1-110-231...	Dec 27, 2023	1	800006	ABC bali 250g	kopi	2	Pcs	14370	28740	28740
10	1-113-231...	Nov 29, 2023	1	800006	ABC bali 250g	kopi	2	Pcs	14370	28740	28740
11	1-19-2309...	Sep 4, 2023	1	800006	ABC bali 250g	kopi	2	Pcs	14370	28740	129740
12	1-15-2307...	Jul 6, 2023	1	800006	ABC bali 250g	kopi	2	Pcs	13710	27420	87070
13	1-15-2306...	Jun 5, 2023	2	800006	ABC bali 250g	kopi	1	Pcs	13710	13710	108860
14	1-17-2307...	Jul 24, 2023	4	800006	ABC bali 250g	kopi	1	Pcs	14130	14130	999960
15	1-15-2307...	Jul 14, 2023	4	800006	ABC bali 250g	kopi	2	Pcs	13710	27420	294180
16	1-15-2306...	Jun 8, 2023	2	800006	ABC bali 250g	kopi	1	Pcs	13710	13710	64790
17	1-17-2307...	Jul 10, 2023	14	800006	ABC bali 250g	kopi	1	Pcs	13710	13710	541270
18	1-19-2307...	Jul 24, 2023	12	800006	ABC bali 250g	kopi	1	Pcs	14130	14130	601160
19	1-110-230...	Jul 25, 2023	7	800006	ABC bali 250g	kopi	2	Pcs	14130	28260	596760
20	1-110-230...	Jul 24, 2023	27	800006	ABC bali 250g	kopi	1	Pcs	14130	14130	1300690

ExampleSet (103,181 examples, 0 special attributes, 11 regular attributes)

Gambar 2. Data Transaksi 1 Tahun

3.2 Cleaning Data

Pembersihan data adalah langkah awal yang krusial dalam pemrosesan data, terutama untuk meningkatkan kualitas data sebelum analisis. Menghapus transaksi dan produk dengan data yang hilang sangat penting untuk dataset yang akan digunakan dalam aplikasi algoritma seperti FP-Growth. Diketahui bahwa sebagian besar waktu digunakan untuk membersihkan dan memproses data, sering kali mencapai 60–80% dari upaya analitis, yang menunjukkan pentingnya menetapkan metodologi pengumpulan dan pembersihan data yang efisien [21], [22]. Teknik otomatis, seperti yang menggunakan algoritma pembelajaran mesin, dapat meningkatkan efisiensi dan presisi pembersihan data dengan mendeteksi dan memperbaiki entri yang hilang atau salah lebih komprehensif daripada metode manual [23]. Selain itu, fokus pada kolom yang relevan, seperti "Nomor Transaksi" dan "Nama Produk," memungkinkan dataset yang lebih ramping yang lebih sesuai dengan tujuan analisis keranjang pasar [8], sehingga meningkatkan kinerja pemodelan secara keseluruhan [8].



Result History: ExampleSet (Aggregate) x

Open in: Turbo Prep, Auto Model, Interactive Analysis

Filter (103,173 / 103,173 examples): all

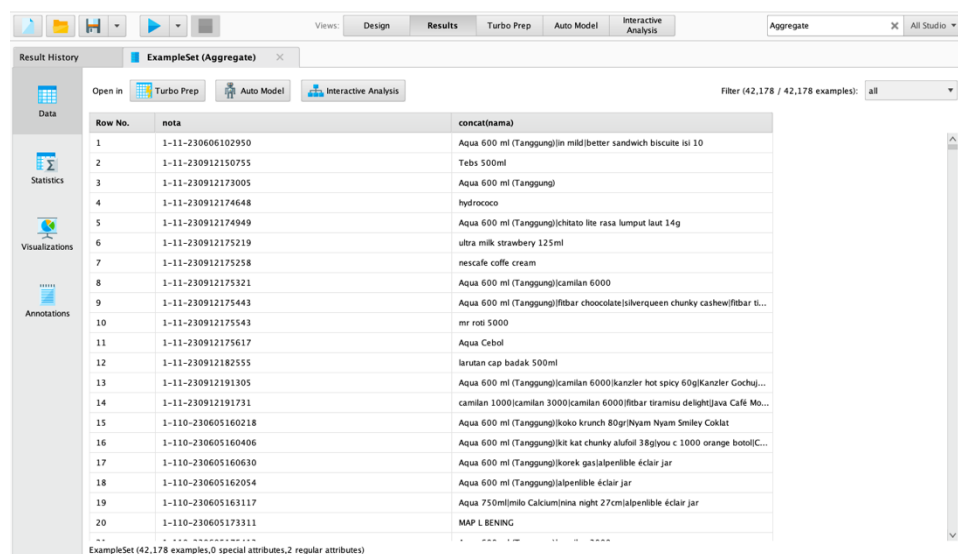
Row No.	nota	nama
1	1-11-230606102950	Aqua 600 ml (Tanggung)
2	1-11-230606102950	better sandwich biscuite isi 10
3	1-11-230606102950	in mild
4	1-11-230912150755	Tebs 500ml
5	1-11-230912173005	Aqua 600 ml (Tanggung)
6	1-11-230912174648	hydrococo
7	1-11-230912174949	Aqua 600 ml (Tanggung)
8	1-11-230912174949	chitato lite rasa lumput laut 14g
9	1-11-230912175219	ultra milk strawberry 125ml
10	1-11-230912175258	nescafe coffe cream
11	1-11-230912175321	Aqua 600 ml (Tanggung)
12	1-11-230912175321	camilan 6000
13	1-11-230912175443	Aqua 600 ml (Tanggung)
14	1-11-230912175443	fibar chocolate
15	1-11-230912175443	fibar tiramisu delight
16	1-11-230912175443	silverqueen chunky cashew
17	1-11-230912175543	mr roti 5000
18	1-11-230912175617	Aqua Cebol
19	1-11-230912182555	larutan cap badak 500ml
20	1-11-230912191305	Aqua 600 ml (Tanggung)

ExampleSet (103,173 examples, 0 special attributes, 2 regular attributes)

Gambar 3. Hasil Cleaning Data Transaksi

3.3 Data Transformation

Setelah melakukan proses pembersihan data, data yang sudah dibersihkan selanjutnya data tersebut akan ditransformasikan untuk meningkatkan performa data mining dengan melakukan aggregation untuk menyatukan data yang memiliki nomor transaksi yang sama. Berikut merupakan hasil dari aggregation data.



Result History: ExampleSet (Aggregate) x

Open in: Turbo Prep, Auto Model, Interactive Analysis

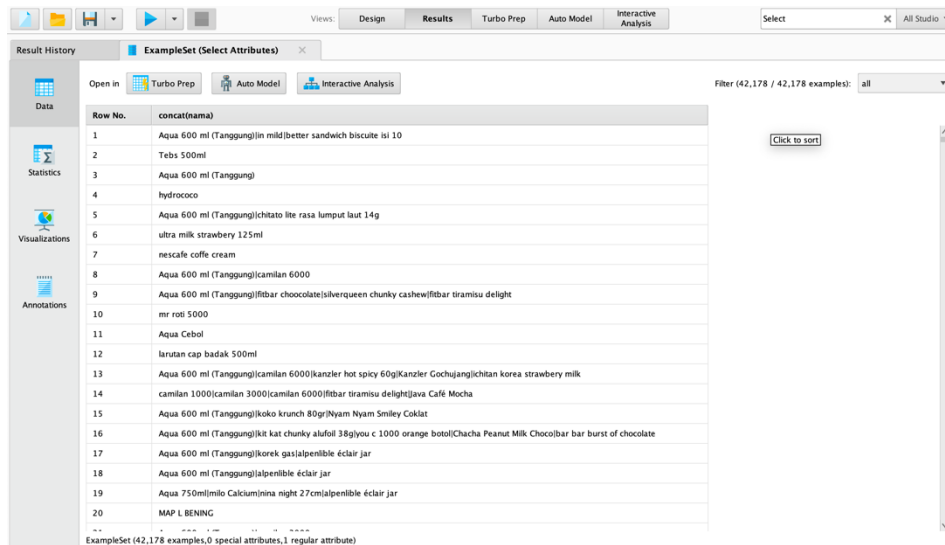
Filter (42,178 / 42,178 examples): all

Row No.	nota	concat(nama)
1	1-11-230606102950	Aqua 600 ml (Tanggung)in mild/better sandwich biscuite isi 10
2	1-11-230912150755	Tebs 500ml
3	1-11-230912173005	Aqua 600 ml (Tanggung)
4	1-11-230912174648	hydrococo
5	1-11-230912174949	Aqua 600 ml (Tanggung)chitato lite rasa lumput laut 14g
6	1-11-230912175219	ultra milk strawberry 125ml
7	1-11-230912175258	nescafe coffe cream
8	1-11-230912175321	Aqua 600 ml (Tanggung)camilan 6000
9	1-11-230912175443	Aqua 600 ml (Tanggung)fibar chocolate/silverqueen chunky cashew/fibar ti...
10	1-11-230912175543	mr roti 5000
11	1-11-230912175617	Aqua Cebol
12	1-11-230912182555	larutan cap badak 500ml
13	1-11-230912191305	Aqua 600 ml (Tanggung)camilan 6000/kanzler hot spicy 60g/kanzler Gochuj...
14	1-11-230912191731	camilan 1000/camilan 3000/camilan 6000/fibar tiramisu delight/java CaFé Mo...
15	1-110-230605160218	Aqua 600 ml (Tanggung)koko krunch 80gr/Nyam Nyam Smiley Coklat
16	1-110-230605160406	Aqua 600 ml (Tanggung)kit kat chunky alufol 38g/you c 1000 orange botol(C...
17	1-110-230605160630	Aqua 600 ml (Tanggung)korek gas/alpenlible éclair jar
18	1-110-230605162054	Aqua 600 ml (Tanggung)alpenlible éclair jar
19	1-110-230605163117	Aqua 750ml/milo Cakum/nina night 27cm/alpenlible éclair jar
20	1-110-230605173311	MAP L BENING

ExampleSet (42,178 examples, 0 special attributes, 2 regular attributes)

Gambar 4. Hasil Aggregation Data

Setelah melakukan aggregation data dengan menggabungkan data nama produk yang memiliki nomor transaksi yang sama dengan memisahkan nama produk tersebut dengan tanda “[” dan jumlah data yang dihasilkan yaitu 42.178 data, tahapan berikutnya yakni memilih satu tipe data yang akan diimplementasikan kedalam algoritma FP-Growth, pada penelitian ini kita bertujuan melihat keterkaitan antar produk jadi kita hanya membutuhkan kolom “Nama Produk” dan kolom “Nomor transaksi” akan dihilangkan sehingga hasilnya akan jadi sebagai berikut.



Row No.	concat(nama)
1	Aqua 600 ml (Tanggung)lin mild(better sandwich biscuite isi 10
2	Tebis 500ml
3	Aqua 600 ml (Tanggung)
4	hydrococo
5	Aqua 600 ml (Tanggung)lchitato lite rasa lumpur laut 14g
6	ultra milk strawberry 125ml
7	nescafe coffe cream
8	Aqua 600 ml (Tanggung)camilan 6000
9	Aqua 600 ml (Tanggung)lftbar chocolate(silverqueen chunky cashewlftbar tiramisu delight
10	mr roti 5000
11	Aqua Cebol
12	larutan cap badak 500ml
13	Aqua 600 ml (Tanggung)camilan 6000(kanzler hot spicy 60g)kanzler Gochujanglchitan korea strawberry milk
14	camilan 1000(camilan 3000)camilan 6000lftbar tiramisu delightjava Cafe Mocha
15	Aqua 600 ml (Tanggung)koko krunch 80gr)Nyam Nyam Smiley Coklat
16	Aqua 600 ml (Tanggung)kit kat chunky alufol 38g)you c 1000 orange botol(Chacha Peanut Milk Chocobar bar burst of chocolate
17	Aqua 600 ml (Tanggung)korek gaslalpenible éclair jar
18	Aqua 600 ml (Tanggung)lalpenible éclair jar
19	Aqua 750ml)mla Calcium)mla night 27cm)lalpenible éclair jar
20	MAP L BENING

Gambar 5. Hasil Select Attribute Aggregation Data

3.4 Implementasi Algoritma FP-Growth

Pada tahapan implementasi algoritma FP-Growth akan dilakukan 2 tahapan yakni membuat frequency table dan pembuatan FP Tree namun dengan menggunakan tools rapidminer atau yang sekarang dikenal sebagai AI Studio maka tahapan implementasi algoritma FP-Growth akan menjadi lebih mudah.

Tahap awal yang perlu diperhatikan yaitu memilih minimum requirement sesuai dengan frequency, pada penelitian ini akan diterapkan minimum frequency yaitu 250 item dan minimum item per itemset yaitu 1. Dimana nilai minimum frequency 250 untuk fokus pada pola yang cukup sering muncul dalam dataset. Dengan ambang batas ini, pola-pola dengan jumlah kemunculan yang signifikan akan diperhatikan, mengurangi gangguan dari pola yang jarang dan kurang relevan. Nilai ini juga dapat memastikan bahwa dataset yang besar tidak menghasilkan pola yang terlalu banyak sehingga sulit dianalisis. Lalu minimum item per itemset ditentukan sebanyak 1 yaitu untuk mengidentifikasi pola dasar, dengan nilai ini dapat ditemukan asosiasi yang relevan dan bermakna antar item dalam dataset.

Parameters
✕

FP-Growth

input format ⓘ

item separators ⓘ

☐ use quotes ⓘ

☒ trim item names ⓘ

min requirement ⓘ

min frequency ⓘ

min items per itemset ⓘ

max items per itemset ⓘ

max number of itemsets ⓘ

☒ find min number of itemsets ⓘ

min number of itemsets ⓘ

must contain list ⓘ

Gambar 6. Filter Data Implementasi Algoritma FP-Growth

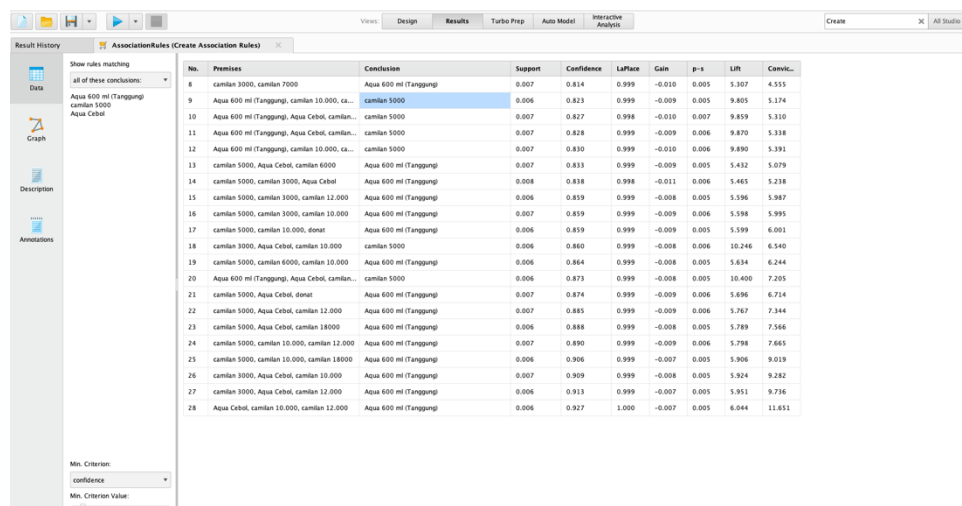
Sehingga akan ditampilkan data sesuai penyaringan data sesuai dengan frequent item yang diinginkan sebagai berikut.

FrequentItemSets (FP-Growth)						
No. of Sets: 183	Size	Support	Item 1	Item 2	Item 3	Item 4
Total Max. Size: 4	1	0.243	aqua tanggung happy dine in			
Min. Size: 1	1	0.133	Aqua 600 ml (Tanggung)			
Max. Size: 4	1	0.084	camilan 10000			
Contains Item:	1	0.073	camilan 10000			
Update View	1	0.060	Aqua Cebol			
	1	0.056	camilan 60000			
	1	0.043	camilan 10.000			
	1	0.042	camilan 12.000			
	1	0.038	AQUA CEROL HAPPY DINE IN			
	1	0.038	donat			
	1	0.032	camilan 18000			
	1	0.027	mikanai special			
	1	0.025	camilan 7000			
	1	0.024	Aqua 750ml			
	1	0.023	camilan 15.000			
	1	0.023	Teh Botol Kosak 330ml			
	1	0.022	mikanai 15 ribu			
	1	0.018	milk 16000			
	1	0.018	tehbotol seora 450ml			
	1	0.018	binge beng caramel 20g			
	1	0.015	fibar chocolate			
	1	0.014	camilan4000			
	1	0.014	fibar multigrain			
	1	0.014	fibar premium delight			
	1	0.014	camilan 10000			
	1	0.013	sari roti sandwich vanilla			
	1	0.010	fibar cheese delight			

Gambar 7. Hasil Frequent Item

Gambar 7 menunjukkan hasil Frequent Itemset Mining menggunakan algoritma FP-Growth dalam sebuah perangkat lunak analisis data. Tampilan ini memperlihatkan daftar kombinasi item yang sering muncul bersama dalam transaksi penjualan berdasarkan nilai support. Support menunjukkan seberapa sering kombinasi item tertentu muncul dalam seluruh transaksi. Pada kolom pertama, terdapat Size, yang menunjukkan jumlah item dalam satu kombinasi. Misalnya, jika size = 1, itu berarti hanya satu item yang sering muncul sendirian, sedangkan size yang lebih besar menunjukkan kombinasi beberapa item yang sering dibeli bersamaan. Kolom berikutnya berisi item-item yang ditemukan dalam kombinasi tersebut, seperti "Aqua 600 ml (Tanggung)", "camilan 5000", "Aqua Cebol", dan berbagai jenis camilan lainnya. Nilai support yang lebih tinggi menunjukkan bahwa item atau kombinasi tersebut sering muncul dalam dataset transaksi. Contohnya, "camilan 3000" memiliki support yang cukup tinggi, menunjukkan bahwa produk ini sering dibeli dalam transaksi yang dianalisis. Selain itu, beberapa kombinasi lain seperti "Aqua 600 ml (Tanggung)" dan "camilan 5000" juga memiliki nilai support yang signifikan, mengindikasikan bahwa mereka sering dibeli bersamaan.

Setelah melakukan penyaringan data selanjutnya akan dicari association rules dengan menyaring nilai confidence dari hasil penerapan algoritma FP-Growth dengan minimum confidence yaitu 80% (0.8) yang menunjukkan hubungan antar dua item berdasarkan suatu kondisi tertentu kemiripan data sehingga dihasilkan association rules yaitu sebagai berikut.



No.	Premises	Conclusion	Support	Confidence	Lift	Gain	p-value	Lift	Conv...
8	camilan 3000, camilan 7000	Aqua 600 ml (Tanggung)	0.007	0.814	0.999	-0.010	0.005	5.307	4.555
9	Aqua 600 ml (Tanggung), camilan 10.000, ca...	camilan 5000	0.006	0.823	0.999	-0.009	0.005	9.805	5.174
10	Aqua 600 ml (Tanggung), Aqua Cebol, camila...	camilan 5000	0.007	0.827	0.998	-0.010	0.007	9.859	5.310
11	Aqua 600 ml (Tanggung), Aqua Cebol, camila...	camilan 5000	0.007	0.828	0.999	-0.009	0.006	9.870	5.338
12	Aqua 600 ml (Tanggung), camilan 10.000, ca...	camilan 5000	0.007	0.830	0.999	-0.010	0.006	9.890	5.391
13	camilan 5000, Aqua Cebol, camilan 6000	Aqua 600 ml (Tanggung)	0.007	0.833	0.999	-0.009	0.005	5.432	5.079
14	camilan 5000, camilan 3000, Aqua Cebol	Aqua 600 ml (Tanggung)	0.008	0.838	0.998	-0.011	0.006	5.465	5.238
15	camilan 5000, camilan 3000, camilan 12.000	Aqua 600 ml (Tanggung)	0.006	0.859	0.999	-0.008	0.005	5.596	5.967
16	camilan 5000, camilan 3000, camilan 10.000	Aqua 600 ml (Tanggung)	0.007	0.859	0.999	-0.009	0.006	5.598	5.995
17	camilan 5000, camilan 10.000, donat	Aqua 600 ml (Tanggung)	0.006	0.859	0.999	-0.009	0.005	5.599	6.001
18	camilan 3000, Aqua Cebol, camilan 10.000	camilan 5000	0.006	0.860	0.999	-0.008	0.006	10.246	6.540
19	camilan 5000, camilan 6000, camilan 10.000	Aqua 600 ml (Tanggung)	0.006	0.864	0.999	-0.008	0.005	5.634	6.244
20	Aqua 600 ml (Tanggung), Aqua Cebol, camila...	camilan 5000	0.006	0.873	0.999	-0.008	0.005	10.400	7.205
21	camilan 5000, Aqua Cebol, donat	Aqua 600 ml (Tanggung)	0.007	0.874	0.999	-0.009	0.006	5.696	6.714
22	camilan 5000, Aqua Cebol, camilan 12.000	Aqua 600 ml (Tanggung)	0.007	0.885	0.999	-0.009	0.006	5.767	7.344
23	camilan 5000, Aqua Cebol, camilan 18000	Aqua 600 ml (Tanggung)	0.006	0.888	0.999	-0.008	0.005	5.789	7.566
24	camilan 5000, camilan 10.000, camilan 12.000	Aqua 600 ml (Tanggung)	0.007	0.890	0.999	-0.009	0.006	5.798	7.665
25	camilan 5000, camilan 10.000, camilan 18000	Aqua 600 ml (Tanggung)	0.006	0.906	0.999	-0.007	0.005	5.906	9.019
26	camilan 3000, Aqua Cebol, camilan 10.000	Aqua 600 ml (Tanggung)	0.007	0.909	0.999	-0.008	0.005	5.924	9.282
27	camilan 3000, Aqua Cebol, camilan 12.000	Aqua 600 ml (Tanggung)	0.006	0.913	0.999	-0.007	0.005	5.951	9.716
28	Aqua Cebol, camilan 10.000, camilan 12.000	Aqua 600 ml (Tanggung)	0.006	0.927	1.000	-0.007	0.005	6.044	11.651

Gambar 8. Association Rules Data Transaksi

Gambar 8 menampilkan hasil Association Rule Mining yang dihasilkan menggunakan algoritma FP-Growth, yang bertujuan untuk menemukan hubungan antara item-item yang sering muncul bersama dalam transaksi penjualan. Tabel pada gambar ini berisi beberapa kolom utama, yaitu Premises, Conclusion, Support, Confidence, dan Lift. Kolom Premises menunjukkan kombinasi item yang sering dibeli bersama dalam transaksi, sementara Conclusion menampilkan item yang kemungkinan besar akan dibeli jika Premises terjadi. Nilai Support menunjukkan seberapa sering kombinasi tersebut muncul dalam seluruh dataset transaksi, sedangkan Confidence menggambarkan seberapa besar kemungkinan seseorang akan membeli Conclusion setelah membeli Premises. Lift memberikan gambaran tentang seberapa kuat hubungan antara Premises dan Conclusion dibandingkan dengan kejadian acak, di mana nilai Lift yang lebih tinggi menunjukkan hubungan yang lebih kuat. Dari tabel ini, terlihat bahwa beberapa produk, seperti Aqua 600 ml (Tanggung) dan berbagai camilan, sering muncul bersama dalam transaksi dengan confidence yang tinggi. Hal ini menunjukkan bahwa jika seorang pelanggan membeli camilan tertentu, mereka memiliki

kemungkinan besar untuk juga membeli Aqua 600 ml (Tanggung). Informasi ini dapat digunakan untuk strategi pemasaran seperti pengaturan tata letak toko untuk meningkatkan kemungkinan pembelian tambahan.

3.5 Hasil Data Mining

Berdasarkan hasil dari penggunaan aplikasi RapidMiner sehingga didapat hasil Data Mining yang digambarkan pada tabel association rules.

Tabel 2. Hasil Data Mining

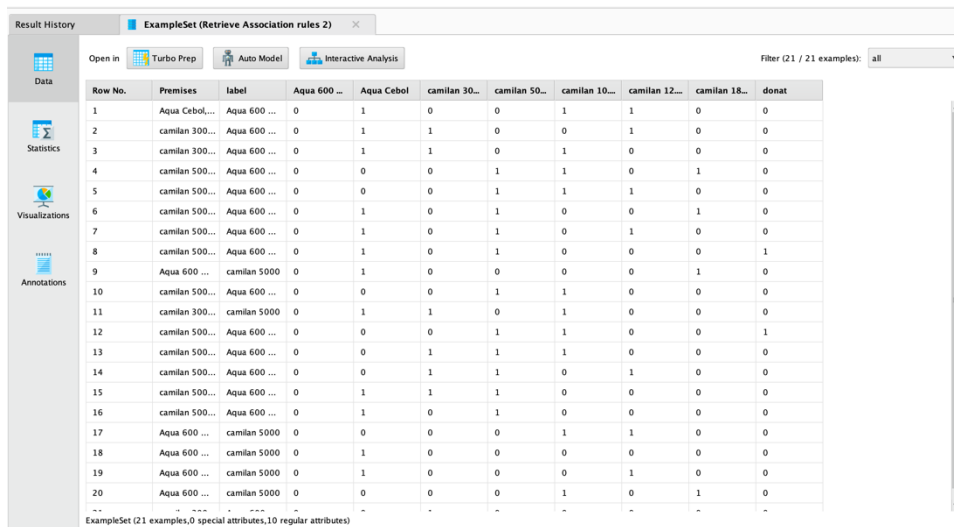
Premises	Conclusion	Support	Confidence	Lift
Aqua Cebol, camilan 10.000, camilan 12.000	Aqua 600 ml (Tanggung)	0.006	0.93	6,0
camilan 3000, Aqua Cebol, camilan 12.000	Aqua 600 ml (Tanggung)	0.006	0.91	5,9
camilan 3000, Aqua Cebol, camilan 10.000	Aqua 600 ml (Tanggung)	0.007	0.91	5,9
camilan 5000, camilan 10.000, camilan 18000	Aqua 600 ml (Tanggung)	0.006	0.91	5,9
camilan 5000, camilan 10.000, camilan 12.000	Aqua 600 ml (Tanggung)	0.007	0.89	5,8
camilan 5000, Aqua Cebol, camilan 18000	Aqua 600 ml (Tanggung)	0.006	0.89	5,8
camilan 5000, Aqua Cebol, camilan 12.000	Aqua 600 ml (Tanggung)	0.007	0.88	5,8
camilan 5000, Aqua Cebol, donat	Aqua 600 ml (Tanggung)	0.007	0.87	5,7
Aqua 600 ml (Tanggung), Aqua Cebol, camilan 18000	camilan 5000	0.006	0.87	10,4
camilan 5000, camilan 6000, camilan 10.000	Aqua 600 ml (Tanggung)	0.006	0.86	5,6
camilan 3000, Aqua Cebol, camilan 10.000	camilan 5000	0.006	0.86	10,2
camilan 5000, camilan 10.000, donat	Aqua 600 ml (Tanggung)	0.006	0.86	5,6
camilan 5000, camilan 3000, camilan 10.000	Aqua 600 ml (Tanggung)	0.007	0.86	5,6
camilan 5000, camilan 3000, camilan 12.000	Aqua 600 ml (Tanggung)	0.006	0.86	5,6
camilan 5000, camilan 3000, Aqua Cebol	Aqua 600 ml (Tanggung)	0.008	0.84	5,5
camilan 5000, Aqua Cebol, camilan 6000	Aqua 600 ml (Tanggung)	0.007	0.83	5,4
Aqua 600 ml (Tanggung), camilan 10.000, camilan 12.000	camilan 5000	0.007	0.83	9,8
Aqua 600 ml (Tanggung), Aqua Cebol, camilan 6000	camilan 5000	0.007	0.83	9,9
Aqua 600 ml (Tanggung), Aqua Cebol, camilan 12.000	camilan 5000	0.007	0.83	9,9
Aqua 600 ml (Tanggung), camilan 10.000, camilan 18000	camilan 5000	0.006	0.82	9,8
camilan 3000, camilan 7000	Aqua 600 ml (Tanggung)	0.007	0.81	5,3

Dari hasil yang didapatkan diatas, peneliti akan melakukan penataan terhadap rak (display) sesuai dengan rekomendasi hasil penggunaan aplikasi RapidMiner untuk menganalisis hasil penjualan sebelum dan sesudah menata kembali rak pada tempat penelitian yakni UNDIKNAS Mart.

3.6 Validasi Data Mining

Sebelum melakukan implementasi hasil dari data mining yang menggunakan algoritma FP-Growth, akan dilakukan sebuah validasi hasil dengan melakukan pengujian menggunakan Cross Validation dan model yang digunakan pada pengujian yaitu decision tree untuk mendapatkan akurasi dari hasil data mining yang sudah dilakukan.

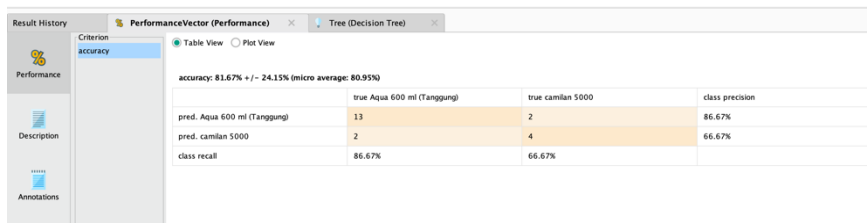
Tahap awal dilakukan format data hasil penerapan algoritma FP-Growth seperti yang ditunjukkan pada tabel 2 dengan data yang berjumlah 21 data diatas menjadi data seperti yang ditunjukkan pada gambar 9 sebagai berikut.



Row No.	Premises	label	Aqua 600 ...	Aqua Cebol	camilan 30...	camilan 50...	camilan 10...	camilan 12...	camilan 18...	donat
1	Aqua Cebol...	Aqua 600 ...	0	1	0	0	1	1	0	0
2	camilan 300...	Aqua 600 ...	0	1	1	0	0	1	0	0
3	camilan 300...	Aqua 600 ...	0	1	1	0	1	0	0	0
4	camilan 500...	Aqua 600 ...	0	0	0	1	1	0	1	0
5	camilan 500...	Aqua 600 ...	0	0	0	1	1	1	0	0
6	camilan 500...	Aqua 600 ...	0	1	0	1	0	0	1	0
7	camilan 500...	Aqua 600 ...	0	1	0	1	0	1	0	0
8	camilan 500...	Aqua 600 ...	0	1	0	1	0	0	0	1
9	Aqua 600 ...	camilan 5000	0	1	0	0	0	0	1	0
10	camilan 500...	Aqua 600 ...	0	0	0	1	1	0	0	0
11	camilan 300...	camilan 5000	0	1	1	0	1	0	0	0
12	camilan 500...	Aqua 600 ...	0	0	0	1	1	0	0	1
13	camilan 500...	Aqua 600 ...	0	0	1	1	1	0	0	0
14	camilan 500...	Aqua 600 ...	0	0	1	1	0	1	0	0
15	camilan 500...	Aqua 600 ...	0	1	1	1	0	0	0	0
16	camilan 500...	Aqua 600 ...	0	1	0	1	0	0	0	0
17	Aqua 600 ...	camilan 5000	0	0	0	0	1	1	0	0
18	Aqua 600 ...	camilan 5000	0	1	0	0	0	0	0	0
19	Aqua 600 ...	camilan 5000	0	1	0	0	0	1	0	0
20	Aqua 600 ...	camilan 5000	0	0	0	0	1	0	1	0
21	Aqua 600 ...	camilan 5000	0	0	0	0	0	0	0	0

Gambar 9 Persiapan Data Validasi Hasil Algoritma FP-Growth

Setelah melakukan format hasil data mining algoritma FP-Growth seperti gambar 9 diatas, dimana formatting data dilakukan dengan menjabarkan semua produk yang memiliki nilai confidence 80%, support melebihi 0.05 dan nilai lift lebih dari satu dari jumlah data 103.181 sebagai parameter pengujian yang sesuai dengan premises. Lalu data tersebut akan diuji atau divalidasi kembali dengan menggunakan schema cross validation dimana didalam pengujian ini akan diterapkan model decision tree untuk menguji sebanyak 10 folds (kelipatan) untuk mencapai akurasi tertinggi dan dengan menerapkan tipe sampling yakni stratified sampling yang mana tipe sampling ini memastikan setiap kelas dalam dataset tetap didalam setiap pengujian yang dilakukan dan mengaju pada data label telah ditetapkan, dimana label ini terdapat 2 buah produk yakni “Aqua 600 ml (Tanggung)” dan “camilan 500” sehingga menghasilkan sebuah confusion metrics seperti yang ditunjukkan pada gambar 10 sebagai berikut.

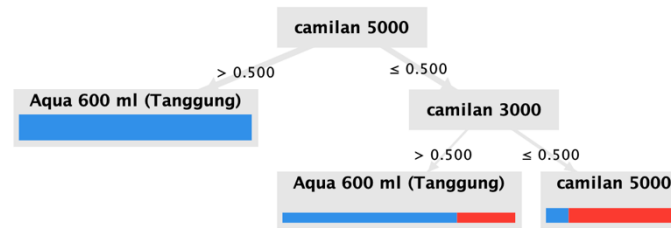


	true Aqua 600 ml (Tanggung)	true camilan 5000	class precision
pred. Aqua 600 ml (Tanggung)	13	2	86.67%
pred. camilan 5000	2	4	66.67%
class recall	86.67%	66.67%	

Gambar 10 Hasil Pengujian CrossValidation Confusion Metrics

Dari gambar 10 diatas dapat dijelaskan hasil pengujian Cross-Validation menggunakan Decision Tree, model memiliki akurasi sebesar $81.67\% \pm 24.15\%$, dengan micro average accuracy sebesar 80.95%. Dari confusion metrics, model memprediksi 13 sampel Aqua 600 ml (Tanggung) dengan benar, tetapi salah memprediksi 2 sampel sebagai camilan 5000. Sementara itu, model berhasil memprediksi 4 sampel camilan 5000 dengan benar, namun salah memprediksi 2 sampel sebagai Aqua 600 ml (Tanggung). Precision untuk kelas Aqua 600 ml (Tanggung) adalah 86.67%, sedangkan untuk camilan 5000 adalah 66.67%. Recall masing-masing kelas juga menunjukkan pola yang serupa, yaitu 86.67% untuk Aqua 600 ml (Tanggung) dan 66.67% untuk camilan 5000. Secara keseluruhan, model cukup baik dalam mengklasifikasikan kedua kategori, meskipun masih terdapat kesalahan terutama pada kelas camilan 5000 yang memiliki recall lebih rendah.

Selain menghasilkan sebuah confusion metrics, pengujian menggunakan cross validation juga menghasilkan sebuah tree dari hasil penerapan model decision tree seperti gambar berikut.

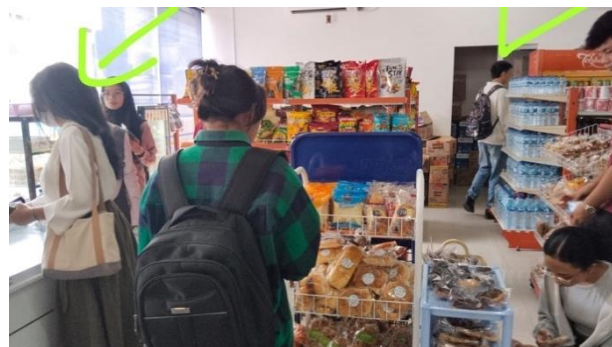


Gambar 11 Hasil Penerapan Model Decision Tree

Pada gambar 11 diatas, hasil penerapan model decision tree menunjukkan struktur pohon keputusan dalam mengklasifikasikan produk berdasarkan fitur tertentu [24]. Node pertama membagi data berdasarkan atribut dengan threshold ≤ 0.500 , di mana jika nilai lebih besar, sistem memprediksi sebagai Aqua 600 ml (Tanggung), sedangkan jika lebih kecil atau sama, data masuk ke kategori camilan 3000. Dari camilan 3000, pohon kembali bercabang dengan threshold baru, memisahkan antara Aqua 600 ml (Tanggung) dan camilan 5000. Visualisasi ini menunjukkan bahwa sebagian besar sampel Aqua 600 ml (Tanggung) diklasifikasikan dengan baik, tetapi terdapat beberapa kesalahan klasifikasi pada cabang yang mengarah ke camilan 5000, yang tampaknya memiliki distribusi kelas yang lebih beragam. Model ini cukup baik dalam memisahkan kategori utama. Sehingga dari hasil pengujian menggunakan cross validation menunjukkan bahwa hasil dari data mining menggunakan algoritma FP-Growth sudah sesuai dan memiliki tingkat akurasi yang tinggi dan siap untuk diimplementasikan kedalam penataan display produk dengan rekomendasi yang telah diberikan sesuai data transaksi 1 tahun terakhir yang dituangkan dalam bentuk association rules pada tabel 2 [25].

3.7 Validasi Data Mining

Setelah hasil data mining didapatkan, rak (display) akan ditata ulang sesuai hasil yang didapatkan. Dimana pada kenyataan yang didapat produk produk seperti Aqua masih berjauhan dengan produk-produk cemilan dan donat seperti yang ditunjukkan pada gambar berikut.



Gambar 12. Tata Letak Produk Sebelum Menerapkan Hasil Data Mining

Dari Gambar 9 dapat dilihat melalui tanda panah yang ada pada gambar, Aqua masih berjarak 1 rak dengan produk camilan dan donat yang merupakan barang konsinyasi dan perlu diubah tata letaknya dengan menempatkan barang barang tersebut pada jarak yang berdekatan sehingga hasilnya akan seperti gambar dibawah ini.



Gambar 13. Tata Letak Produk Setelah Menerapkan Hasil Data Mining

Pada Gambar 10 dapat dilihat produk – produk yang berkaitan sudah diletakkan secara berdekatan sehingga harapannya dapat meningkatkan hasil penjualan pada UNDIKNAS Mart. Hasil penjualan selama 4 bulan akan dianalisis dimulai dari bulan Oktober 2024 hingga Januari 2025. Akan dilakukan perbandingan terhadap pendapatan, persentase pendapatan, item terjual dan persentase item terjual dari setiap produk yang dianalisis sesuai hasil algoritma FP-Growth. Berikut merupakan perbandingan data penjualan sebelum dan sesudah menerapkan data mining.

Tabel 3. Perbandingan Pendapatan Oktober 2024 Sampai Januari 2025

Produk	Aqua 600 ml (Tanggung)	Aqua Cebol	Camilan 3000	Camilan 5000	Camilan 10.000	Camilan 12.000	Camilan 18000	Donat
Pendapatan Oktober 2024	Rp. 14.240.000	Rp3,201,000	Rp7,497,000	Rp11,227,410	Rp15,830,000	Rp11,964,000	Rp7,524,000	Rp2,982,000
Pendapatan Novemembr 2024	Rp9,984,000	Rp2,280,000	Rp4,170,790	Rp6,027,030	Rp9,550,000	Rp6,936,000	Rp5,706,000	Rp2,009,000
Pendapatan Desember 2024	Rp10,048,000	Rp2,640,000	Rp4,365,000	Rp6,552,500	Rp10,700,000	Rp10,905,890	Rp4,371,100	Rp1,666,000
Pendapatan Januari 2025	Rp2,052,000	Rp393,000	Rp438,000	Rp951,900	Rp400,000	Rp360,000	Rp73,480	Rp0

Tabel 4. Perbandingan Persentase Pendapatan Oktober 2024 Sampai Januari 2025

Produk	Aqua 600 ml (Tanggung)	Aqua Cebol	Camilan 3000	Camilan 5000	Camilan 10.000	Camilan 12.000	Camilan 18000	Donat
Persentase Pendapatan Oktober 2024	3.91%	0.88%	2.06%	3.08%	4.34%	3.28%	2.06%	0.82%
Persentase Pendapatan Novemembr 2024	4.77%	1.09%	1.99%	2.88%	4.56%	3.31%	2.72%	0.96%
Persentase Pendapatan Desember 2024	4.51%	1.19%	1.96%	2.94%	4.81%	4.9%	1.96%	0.75%
Pendapatan Januari 2025	3.49%	0.67%	0.75%	1.62%	0.68%	0.61%	0.13%	0%

Tabel 5. Perbandingan Item Terjual Oktober 2024 Sampai Januari 2025

Produk	Aqua 600 ml (Tanggung)	Aqua Cebol	camilan 3000	camilan 5000	camilan 10.000	camilan 12.000	camilan 18000	donat
Item Terjual Oktober 2024	1454	615	887	960	779	542	234	135

Item Terjual Novemembr 2024	961	404	527	506	435	285	156	103
Item Terjual Desember 2024	1194	562	634	667	514	494	139	92
Item Terjual Januari 2025	254	75	85	123	31	23	3	0

Tabel 6. Perbandingan Persentase Item Terjual Oktober 2024 Sampai Januari 2025

Produk	Aqua 600 ml (Tanggung)	Aqua Cebol	camilan 3000	camilan 5000	camilan 10.000	camilan 12.000	camilan 18000	donat
Persentase Item Terjual Oktober 2024	6.66%	2.82%	4.07%	4.4%	3.57%	2.48%	1.07%	0.62%
Persentase Item Terjual Novemembr 2024	7.68%	3.23%	4.21%	4.04%	3.47%	2.28%	1.25%	0.82%
Persentase Item Terjual Desember 2024	8.13%	3.83%	4.32%	4.54%	3.5%	3.36%	0.95%	0.63%
Persentase Item Terjual Januari 2025	7.32%	2.16%	2.45%	3.54%	0.89%	0.66%	0.09%	0%

Dari tabel perbandingan pendapatan, persentase pendapatan, item terjual serta persentase item terjual. Pendapatan di bulan November hingga Januari mengalami penurunan dibandingkan dengan bulan Oktober. Hal ini disebabkan oleh adanya pembangunan yang menutupi akses ke Undiknas Mart. Meskipun pendapatan secara nominal menurun, persentase item terjual mengalami peningkatan dan lebih populer dari sebelum dilakukannya data mining.

4. Kesimpulan

Dalam Berdasarkan hasil penelitian yang dilakukan di UNDIKNAS Mart, penerapan data mining dengan menggunakan algoritma FP-Growth telah terbukti efektif dalam mengidentifikasi pola pembelian konsumen dan membantu menentukan strategi penataan produk yang optimal. Proses dimulai dengan pengumpulan data transaksi penjualan selama satu tahun, yang kemudian melalui tahap pembersihan (data cleaning) untuk memastikan bahwa hanya data yang relevan, seperti nomor transaksi dan nama barang, yang digunakan dalam analisis. Selanjutnya, data yang telah dibersihkan dilakukan transformasi melalui agregasi untuk mempersiapkannya dalam penerapan algoritma FP-Growth. Algoritma ini digunakan untuk menghitung nilai support, confidence, dan lift yang menjadi indikator utama dalam analisis pola pembelian.

Hasil dari penerapan algoritma FP-Growth, yang menggunakan tools RapidMiner, menunjukkan bahwa dengan jumlah data transaksi sebanyak 103.181 transaksi, nilai support yang diperoleh berkisar antara 0,6% hingga 0,8%. Meskipun terlihat kecil, nilai ini cukup signifikan dalam konteks volume data yang besar, memberikan informasi yang berguna untuk menentukan pola pembelian yang sering terjadi. Nilai confidence yang diperoleh berkisar antara 81% hingga 93%, yang menunjukkan kekuatan asosiasi yang cukup tinggi antara produk-produk yang sering dibeli bersama. Selain itu, nilai lift yang lebih besar dari 1 mengindikasikan hubungan yang kuat antara produk dalam transaksi nyata, bukan sekadar kebetulan.

Pengujian lebih lanjut dengan menggunakan metode cross-validation pada model decision tree menghasilkan akurasi sebesar 81,67%, yang menunjukkan bahwa rekomendasi yang dihasilkan oleh algoritma FP-Growth dapat dipercaya dan relevan. Berdasarkan hasil tersebut, penataan tata letak produk (display) di UNDIKNAS Mart telah diterapkan dengan memperhatikan pola pembelian yang ditemukan dari analisis data mining. Hasilnya menunjukkan adanya peningkatan penjualan yang signifikan, baik dari segi pendapatan maupun item yang terjual, yang tercatat selama periode November hingga Januari setelah implementasi algoritma FP-Growth.

Secara keseluruhan, dapat disimpulkan bahwa penerapan data mining dengan algoritma FP-Growth berhasil meningkatkan penjualan dengan memberikan rekomendasi yang tepat terkait kombinasi produk yang sebaiknya dipajang bersama, serta tata letak produk yang lebih efektif dalam menarik perhatian pelanggan.

Daftar Pustaka

- [1] N. Sahari, A. A. Othman, N. M. Ali, A. Djajanti, and N. Kamarolzaman, "Unravelling the Impact of Visual Merchandising on Consumer Impulse Buying Behaviour in Hypermarket," *Environment-Behaviour Proceedings Journal*, vol. 9, no. 28, pp. 39–45, Mar. 2024, doi: 10.21834/e-bpj.v9i28.5874.
- [2] A. Jaini, W. F. Wan Zulkiffli, M. Ismail, F. Mohd, and H. Hussin, "Understanding the Influence of Visual Merchandising on Consumers' Impulse Buying Behaviour: Analysing the Literature," *International Journal of Academic Research in Business and Social Sciences*, vol. 11, no. 12, Dec. 2021, doi: 10.6007/IJARBS/v11-i12/11277.
- [3] I. Lutfiani, H. Farisi, and A. Yuliana, "Pengaruh Hedonic Shopping Motivation, Visual Merchandising, dan Store Atmosphere Terhadap Impulse Buying Melalui Positive Emotion Sebagai Variabel Intervening Pada Konsumen Industri Retail Tiongkok di Indonesia," *JURNAL EKONOMI, MANAJEMEN, BISNIS, DAN SOSIAL (EMBISS)*, vol. 4, no. 1, pp. 74–88, Nov. 2023, doi: 10.59889/embiss.v4i1.282.
- [4] J. Chen and Y.-W. Chang, "How smart technology empowers consumers in smart retail stores? The perspective of technology readiness and situational factors," *Electronic Markets*, vol. 33, no. 1, p. 1, Dec. 2023, doi: 10.1007/s12525-023-00635-6.
- [5] K. Czerniachowska, "Merchandising rules for shelf space allocation with horizontal and vertical positions," *Informatyka Ekonomiczna*, vol. 2021, no. 4, pp. 9–33, 2021, doi: 10.15611/ie.2021.1.01.
- [6] M. C. Cant and C. Bothma, "Fashion retail strategies in-store design and planning: the case of South Africa," *Entrepreneurship and Sustainability Issues*, vol. 10, no. 4, pp. 408–427, Jun. 2023, doi: 10.9770/jesi.2023.10.4(25).
- [7] S. Rana and M. N. I. Mondal, "Seasonal and Multilevel Association Based Approach for Market Basket Analysis in Retail Supermarket," *European Journal of Information Technologies and Computer Science*, vol. 1, no. 4, pp. 9–15, Oct. 2021, doi: 10.24018/compute.2021.1.4.31.
- [8] H. Ghous, M. Malik, and I. Rehman, "Deep Learning based Market Basket Analysis using Association Rules," *KIET Journal of Computing and Information Sciences*, vol. 6, no. 2, pp. 14–34, Aug. 2023, doi: 10.51153/kjicis.v6i2.166.
- [9] S. A. Rahayu, "Market Basket Analysis Using FP-Growth Algorithm to Design Marketing Strategy by Determining Consumer Purchasing Patterns," *Journal of Applied Data Sciences*, vol. 4, no. 1, pp. 38–49, Jan. 2023, doi: 10.47738/jads.v4i1.83.
- [10] I. G. T. Isa, "Sales Strategy by Determining Tenant Position at Sriwijaya State Polytechnic Through Association Rules Analysis," *IJEED (International Journal of Entrepreneurship and Business Development)*, vol. 5, no. 3, pp. 563–573, May 2022, doi: 10.29138/ijeed.v5i3.1855.
- [11] T. Han, W. Wang, M. Guo, and S. Ning, "Association Rules Mining Algorithm Based on Information Gain Ratio Attribute Reduction," 2022, pp. 181–189. doi: 10.1007/978-3-030-92632-8_18.
- [12] A. Boroumand, S. Ghose, G. F. Oliveira, and O. Mutlu, "Polynesia: Enabling Effective Hybrid Transactional/Analytical Databases with Specialized Hardware/Software Co-Design," Mar. 2021, [Online]. Available: <http://arxiv.org/abs/2103.00798>
- [13] H. Henderi, R. Irawatia, I. Indra, D. A. Dewi, and T. B. Kurniawan, "Big Data Analysis using Elasticsearch and Kibana: A Rating Correlation to Sustainable Sales of Electronic Goods," *HighTech and Innovation Journal*, vol. 4, no. 3, pp. 583–591, Sep. 2023, doi: 10.28991/HIJ-2023-04-03-09.
- [14] L. Samboteng, R. Rulinawaty, K. Rachmat, M. Basit, and R. Rahim, "Market basket analysis of administrative patterns data of consumer purchases using data mining technology," *Journal of Applied Engineering Science*, vol. 20, no. 2, pp. 339–345, 2022, doi: 10.5937/jaes0-32019.



- [15] D. Hartanti and V. Atina, "Product Stock Supply Analysis System with FP Growth Algorithm," *Journal of Information Systems and Informatics*, vol. 5, no. 4, pp. 1312–1320, Dec. 2023, doi: 10.51519/journalisi.v5i4.580.
- [16] M. Atif, "Data mining," *International Journal of Communication and Information Technology*, vol. 3, no. 1, pp. 37–40, Jan. 2022, doi: 10.33545/2707661X.2022.v3.i1a.44.
- [17] A. N. Habyba, R. Fitriana, and T. Theodora, "Quality Improvement of NH1X36B Pre-Printed Box with QM-CRISP DM Approach at PT X," *Operations Excellence Journal of Applied Industrial Engineering*, vol. 13, no. 3, p. 298, Nov. 2021, doi: 10.22441/oe.2021.v13.i3.028.
- [18] J. Pacheco-Ortiz, L. Rodríguez-Mazahua, J. Mejía-Miranda, I. Machorro-Cano, and U. Juárez-Martínez, "Towards Association Rule-Based Item Selection Strategy in Computerized Adaptive Testing," 2021, pp. 27–54. doi: 10.1007/978-3-030-71115-3_2.
- [19] F. M. Raihan and Y. Miftahuddin, "Penerapan Algoritma Apriori Pada Riwayat Data Kecelakaan Lalu Lintas," *Infotek: Jurnal Informatika dan Teknologi*, vol. 5, no. 1, pp. 62–71, Jan. 2022, doi: 10.29408/jit.v5i1.4402.
- [20] T. Uçar and A. Karahoca, "Benchmarking data mining approaches for traveler segmentation," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 1, p. 409, Feb. 2021, doi: 10.11591/ijece.v11i1.pp409-415.
- [21] N. S. Chauhan, "Data Cleaning: Challenges and Existing Solutions," *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, vol. 08, no. 04, pp. 1–5, Apr. 2024, doi: 10.55041/IJSREM30377.
- [22] R. R. Balise *et al.*, "Data cleaning and harmonization of clinical trial data: Medication-assisted treatment for opioid use disorder," *PLoS One*, vol. 19, no. 11, p. e0312695, Nov. 2024, doi: 10.1371/journal.pone.0312695.
- [23] M. A. Oladipupo, P. C. Obuzor, B. J. Bamgbade, A. E. Adeniyi, K. M. Olagunju, and S. A. Ajagbe, "An Automated Python Script for Data Cleaning and Labeling using Machine Learning Technique," *Informatica*, vol. 47, no. 6, Jun. 2023, doi: 10.31449/inf.v47i6.4474.
- [24] K. N. Babar, "Performance Evaluation of Decision Trees with Machine Learning Algorithm," *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, vol. 08, no. 05, pp. 1–5, May 2024, doi: 10.55041/IJSREM34179.
- [25] D. Hartanti and V. Atina, "FP GROWTH ALGORITHM MODELING FOR PRODUCT INVENTORY ANALYSIS," *Proceeding of International Conference on Science, Health, And Technology*, pp. pp. 178–184, Sep. 2023, doi: 10.47701/icohetech.v4i1.3390.