

# Implementasi Data Mining Pada Penjualan Buku Dengan Algoritma FP-GROWTH Menggunakan Rapidminer

Zaehol Fatah<sup>1</sup>, Muhammad Al Madany<sup>2</sup>  
<sup>1,2</sup> Universitas Ibrahimy, Situbondo  
Email: muhammadalmadany73@gmail.com

## Abstrak

Studi ini menerapkan algoritma FP-Growth di RapidMiner untuk mengidentifikasi pola pembelian dalam dataset penjualan buku Kaggle. Menggunakan proses KDD, atribut seperti genre dan penerbit ditransformasikan menjadi itemset transaksional untuk analisis asosiasi. Hasilnya menunjukkan bahwa genre fiksi paling sering muncul, dan Penguin Group (USA) LLC memiliki korelasi yang kuat dengan genre ini, ditunjukkan dengan nilai keyakinan 83%. Temuan ini dapat mendukung strategi pemasaran yang tepat sasaran dan keputusan rekomendasi produk.

**Kata Kunci:** FP-Growth, Data Mining, Penjualan Buku, RapidMiner, KDD

## Abstract

This study applies the FP-Growth algorithm in RapidMiner to identify purchasing patterns in a Kaggle book sales dataset. Using the KDD process, attributes such as genre and publisher were transformed into transactional itemsets for association analysis. The results show that the *fiction* genre appears most frequently, and Penguin Group (USA) LLC has a strong correlation with this genre, indicated by an 83% confidence value. These findings can support targeted marketing strategies and product recommendation decisions.

**Keywords:** FP-Growth, Data Mining, Book Sales, RapidMiner, KDD

## PENDAHULUAN

Perkembangan teknologi informasi yang berlangsung sangat cepat telah membawa perubahan besar dalam cara data dikelola dan dimanfaatkan, khususnya pada sektor perdagangan dan penjualan. Setiap transaksi yang terjadi menghasilkan data yang tersimpan dalam basis data perusahaan, namun sering kali belum dianalisis secara optimal untuk mendukung proses pengambilan keputusan. Disinilah konsep data mining menjadi sangat penting, bagi keberlangsungan keputusan bisnis. Data mining berfungsi mengekstraksi pola serta pengetahuan berharga dari himpunan data berukuran besar untuk mendukung pengambilan keputusan. (Han, Kamber, & Pei, 2012). Dalam dunia bisnis, memahami kinerja Beberapa konsep fundamental melekat pada penerapan algoritma asosiasi,

yang secara khusus dapat dijumpai pada algoritma FP-Growth. (Sari & Saputra, 2021).

Yang melalui penerapan algoritma asosiasi, ada beberapa konsep yang di algoritmat asosiasi khususnya algoritma FP-Growth (Frequent pattern growth), Algoritma FP-Growth berfungsi untuk mengidentifikasi itemset yang sering muncul dalam transaksi tanpa melalui proses pembuatan kandidat seperti pada metode Apriori. Teknik ini menyusun struktur pohon yang merepresentasikan frekuensi kemunculan item sehingga proses pencariannya menjadi lebih efisien. [Han, Kamber, & Pei, 2012].

Kajian ini dimaksudkan untuk mengimplementasikan metode terhadap data transaksi penjualan buku buku guna dalam rangka mengidentifikasi

keterhubungan antarproduk yang sering dibeli bersamaan. Hasil yang diperoleh berupa pola pembelian pelanggan yang dapat dimanfaatkan untuk mengetahui produk mana saja yang memiliki kecenderungan dijual atau dibeli dalam kombinasi tertentu (Prasetyo, 2020).

Dalam penerapannya, penelitian ini memanfaatkan perangkat lunak RapidMiner karena memiliki antarmuka visual yang memudahkan proses pemodelan data mining. RapidMiner mendukung berbagai algoritma analisis data dan telah banyak digunakan dalam penelitian dan pengembangan model data mining karena kemudahan penggunaan serta fiturnya yang lengkap (Vulandari, 2017; UTDI, n.d.).

### METODE PENELITIAN

Pada penelitian terdahulu dilakukan oleh Prasetyo, F., & Hasugian, H. ini muncul dari masalah nyata yang dihadapi oleh PT Kusno Baso Ciledug, sebuah perusahaan yang bergerak di bidang produk olahan bakso dan otak-otak. Perusahaan tersebut mengalami penurunan penjualan produk tunai pabrik selama periode dua tahun (Mei 2021 hingga Mei 2023), yang pada akhirnya menyebabkan penumpukan stok di gudang. Masalah ini diyakini terjadi karena pihak manajemen kurang memiliki pengetahuan yang mendalam mengenai produk-produk dengan kombinasi pembelian tertinggi dalam satu transaksi dalam satu transaksi. Studi tersebut dilakukan untuk menganalisis hubungan antarproduk menggunakan metode FP-Growth sebagai dasar penyusunan strategi penjualan silang serta peningkatan performa penjualan perusahaan. Metode yang digunakan adalah Association Rule Mining dengan kerangka CRISP-DM meliputi enam tahap mulai dari pemahaman bisnis hingga penerapan. Data yang dianalisis terdiri dari 600 transaksi tunai pabrik selama dua tahun. Proses pemodelan diterapkan menggunakan perangkat lunak RapidMiner dengan ambang batas dukungan dan kepercayaan yang telah

ditentukan untuk menghasilkan pola hubungan antarproduk.. Hasilnya diperoleh 9 aturan asosiasi, dengan pola terkuat yaitu pembelian *Bakso Kecil 500g* sering diikuti *Bakso Urat 500g*. Nilai support 31,3%, confidence 70,9%, dan lift ratio 1,453 menunjukkan korelasi positif yang kuat. Temuan ini dapat diterapkan untuk strategi cross-selling, promosi bundling, dan penataan produk agar meningkatkan penjualan[Prasetyo & Hasugian, 2023]

CV Jakarta Powder Drink menghadapi tantangan dalam memahami minat konsumen terhadap berbagai varian rasa produknya. Data transaksi yang besar hanya digunakan sebagai arsip tanpa analisis mendalam. Oleh karena itu, dibutuhkan penerapan data mining untuk menggali pola pembelian yang tersembunyi. Pendekatan association rule mining dapat membantu mengidentifikasi hubungan antarproduk. Dengan analisis ini, perusahaan dapat menentukan dimanfaatkan untuk mengidentifikasi kecenderungan produk yang kerap dibeli secara bersamaan penelitian terdiri dari 504 transaksi penjualan mencakup 52 kategori rasa. Proses *preprocessing* dilakukan dengan pembersihan dan pengkodean data menggunakan Jupyter Notebook (Python). Teknik FP-Growth diaplikasikan untuk mengekstraksi itemset yang sering terjadi serta menghasilkan association rule. Perhitungan menggunakan rumus:

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

$$Confidence(A \rightarrow B) = \frac{\text{Transaksi mengandung A dan B}}{\text{Transaksi mengandung A}}$$

Evaluasi dilakukan dengan parameter lift, leverage, dan conviction, menghasilkan 8 aturan asosiasi yang menggambarkan minat konsumen terhadap kombinasi rasa tertentu[Andika, 2022].

Penelitian ini dilatarbelakangi oleh meningkatnya volume data transaksi pada usaha Es Teh Solo Jatibarang yang belum dimanfaatkan secara optimal untuk strategi penjualan dan pengelolaan stok. Dengan memanfaatkan teknologi data mining, penelitian ini bertujuan menemukan

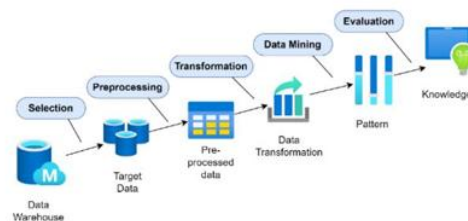
keterkaitan antarproduk guna meningkatkan efisiensi bisnis ritel. Metodologi Proses penelitian mengadopsi metodologi Knowledge Discovery in Database (KDD) melalui lima tahap kunci, mulai dari pemilihan data hingga evaluasi. Proses analisis dilakukan menggunakan perangkat lunak RapidMiner dengan penerapan algoritma FP-Growth untuk menghasilkan *association rules*. Rumus yang digunakan meliputi support, Output dari penerapan algoritma tersebut berfungsi sebagai fondasi dalam perancangan strategi bundling dan tata letak produk guna meningkatkan efektivitas penjualan antarproduk. Hasil dari penerapan algoritma ini menjadi dasar perumusan strategi bundling dan penataan produk agar penjualan lebih optimal (Pratama, 2020).

## HASIL DAN PEMBAHASAN

Riset ini memanfaatkan pendekatan kuantitatif berjenis deskriptif guna menjelaskan objek penelitian secara sistematis berdasarkan fenomena yang diamati. sebagai dasar untuk menggambarkan objek penelitian secara sistematis sesuai fenomena yang diamati (Sugiyono, 2017).

Sumber data berasal dari dataset penjualan buku yang tersedia di platform kaggle. Analisis data dilakukan dengan mengikuti tahapan Data dianalisis menggunakan rangkaian langkah dalam KDD, dan algoritma FP-Growth dimanfaatkan untuk mendeteksi pola asosiasi pada data transaksi. [Kaggle, 2020]. merupakan serangkaian tahapan yang komprehensif memperoleh informasi yang bermanfaat dari data. Proses KDD mencakup beberapa langkah berurutan, dimulai dari pemilihan data yang relevan, pembersihan data dari kesalahan, transformasi ke format yang sesuai, penerapan metode penambangan data, hingga evaluasi pola yang muncul untuk memastikan hasilnya bermanfaat. (Han, Kamber, & Pei, 2012). Metode Knowledge Discovery in Database (KDD) diimplementasikan melalui serangkaian

tahapan penelitian yang terstruktur dan sistematis, dimulai dari proses seleksi data, pra-pemrosesan, transformasi, eksplorasi pola melalui teknik proses penambangan data, sampai pada tahap menilai hasilnya guna mendapatkan informasi yang bermakna. bermakna dari kumpulan data.



Gambar 1. KDD (Knowledge Discovery in Database)

### A. Metode Pengambilan Data

Penelitian ini memanfaatkan metode pengambilan termasuk dalam kategori data sekunder. Data tersebut didapatkan melalui sumber yang sudah tersedia, seperti dokumen, arsip, publikasi, maupun dataset yang telah dikumpulkan sebelumnya [Sugiyono, 2017]. Dataset penelitian diunduh langsung dari Kaggle, yang menyediakan data dalam format terstruktur seperti CSV, sehingga memudahkan proses analisis.

Penggunaan data dari kaggle dipilih karena dataset bersifat terbuka dan gratis telah diverifikasi dan digunakan secara luas serta memiliki ukuran data yang cukup besar variabel yang sesuai untuk proses analisis KDD demikian juga metode pengambilan data sekunder ini juga menjadi dasar awal dalam tahapan seleksi data proses KDD.

#### 1. Tahapan proses KDD (Knowledge Discovery in Database)

- a. Seleksi Data (Data Selection)  
Pemilihan data adalah proses mengambil data yang relevan dari basis data untuk proses KDD (Fayyad et al., 1996).
- b. Praproses data (Preprocessing)  
Prapemrosesan data meliputi pembersihan, integrasi, transformasi, dan reduksi untuk

- mempersiapkan data untuk pengembangan (Rahman, 2019).
- c. Transformasi Data(Transformatino) Tahap transformasi dilakukan dengan mengubah format data agar dapat digunakan untuk proses penambangan, misalnya melalui proses normalisasi atau penggabungan nilai tertentu (Han, Kamber, & Pei, 2012).
- d. Data mining Penambangan data adalah inti dari proses KDD, di mana teknik cerdas digunakan untuk memperoleh pola dari informasi tersebut (Fayyad et al., 1996).
- e. Evaluasi dan Interpretasi Pola Evaluasi pola yang ditambah diperlukan penilaian terhadap pola yang dihasilkan dibutuhkan guna menentukan wawasan yang benar-benar relevan serta menafsirkan temuan tersebutdan menafsirkan hasilnya (Nugroho, 2021).

**B. Dataset**

Penelitian ini memanfaatkan dataset yang berasal dari kumpulan data buku yang memuat beragam informasi mengenai karakteristik, penilaian, serta kinerja penjualan masing-masing buku. Dataset tersebut terdiri atas sejumlah atribut utama Antara lain:

1. publishing year – tahun terbit buku, menunjukkan periode waktu publikasi
2. book name – judul buku yang menjadi objek analisis.
3. Author – nama penulis buku.
4. Language code – kode Bahasa yang digunakan dalam buku
5. Author rating – kategori Tingkat reputasi atau penelitian terhadap penulis
6. Book average rating – nilai rata-rata penilaian buku berdasarkan ulasan pembaca
7. Book ratings count – jumlah total penilaian yang diberikan oleh pembaca
8. Genre – kategori atau jenis buku

9. Gross sales – total pendapatan kotor dari penjualan buku
10. Publisher revenue- pendapatan bersih yang diterima oleh penerbit
11. Sale price – harga jual per eksemplar buku
12. Sales rank – peringkat penjualan buku berdasarkan performa pasar
13. Publisher – nama penerbit buku
14. Units sold – jumlah unit buku yang terjual

Dataset ini terdiri dari sejumlah data yang mewakili berbagai genre dan penerbit besar seperti harpercollins publishers, amazom digital services, inc., Hachette book group, dan penguin group (usa) LCC.

publishing year	book name	author	language code	author rating	book average rating	book ratings count	gross sales	publisher revenue	sale price	units sold	sales rank	publisher
2015	Beowulf	Unknown, Unknown, Unknown	en-US	None	4.74	10000	genre Total: 24000	24000	0.00	0	0	HarperCollins Publishers, Inc.
2017	How One	Frank Wright, Scott McLaughlin, RichardLukas, Dennis	en-US	intermediate	4.25	10000	genre Total: 12475	12475	124.75	1.00	0	HarperCollins Publishers, Inc.
2015	So Good It Hurts	Harper Lee	en-US	None	3.75	10000	genre Total: 4775	3877	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 4225	2475	0.00	0	0	Amazon Digital Services, Inc.
2015	Diagrams of Mind & Sense	Liam Taylor	en-US	intermediate	4.25	10000	genre Total: 3762.5	2272.5	0.00	0	0	Amazon Digital Services, Inc.
2012	The Curious	Victoria Wood	en-US	intermediate	4.25	10000	genre Total: 3950	0	0.00	0	0	Amazon Digital Services, Inc.
2013	The Power of Habit	Charles Duhigg	en-US	intermediate	4.49	10000	genre Total: 2540.07	2448.00	0.00	0	0	HarperCollins Publishers, Inc.
2016	History of the Children of Israel and of the People of the Kingdom of Israel	John Barclay	en-US	intermediate	4.25	10000	genre Total: 2050	0	0.00	0	0	Amazon Digital Services, Inc.
2012	Project	Colleen Hoover	en-US	intermediate	4.34	10000	genre Total: 2000.07	1800.00	0.00	0	0	HarperCollins Publishers, Inc.
2016	The House on the Edge of the World	Patricia Highsmith, Nancy Bond	en-US	intermediate	4.25	10000	genre Total: 1975.50	1875.00	0.00	0	0	Amazon Digital Services, Inc.
2016	The York House	Sean Wilentz	en-US	intermediate	4.11	17962	genre Total: 1794	0	0.00	0	0	Amazon Digital Services, Inc.
2016	The House on the Edge of the World	Patricia Highsmith, Nancy Bond	en-US	intermediate	4.25	10000	genre Total: 1762.50	1662.50	0.00	0	0	Amazon Digital Services, Inc.
2012	Janet Keeler	Barbara Kingsolver	en-US	intermediate	4.42	24076	genre Total: 1642.34	0	0.00	0	0	Amazon Digital Services, Inc.
2016	The House on the Edge of the World	Patricia Highsmith, Nancy Bond	en-US	intermediate	4.25	10000	genre Total: 1612.50	1512.50	0.00	0	0	Amazon Digital Services, Inc.
2012	The House on the Edge of the World	Patricia Highsmith, Nancy Bond	en-US	intermediate	4.25	10000	genre Total: 1512.50	1412.50	0.00	0	0	Amazon Digital Services, Inc.
2017	So Good It Hurts	Harper Lee	en-US	intermediate	3.75	10000	genre Total: 1427.50	0	0.00	0	0	Amazon Digital Services, Inc.
2017	The House on the Edge of the World	Patricia Highsmith, Nancy Bond	en-US	intermediate	4.25	10000	genre Total: 1362.50	1262.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joyner	en-US	intermediate	4.25	10000	genre Total: 1262.50	1162.50	0.00	0	0	Amazon Digital Services, Inc.
2016	When You're in the Middle of a Storm	David Joy										

Row No.	Index	Publishing ...	Book Name	Author	language_c...	Author_Rati...	Book_evera...	Book_rati
1	0	1975	Beowulf	Unknown, Se...	en-US	Novice	3.420	155903
2	1	1987	Batman: Year...	Frank Miller, ...	eng	Intermediate	4.230	145267
3	2	2015	Go Set a Wat...	Harper Lee	eng	Novice	3.310	138669
4	3	2008	When You Ar...	David Sedaris	en-US	Intermediate	4.040	150998
5	4	2011	Daughter of ...	Laini Taylor	eng	Intermediate	4.040	198283
6	5	2015	Red Queen	Victoria Avey...	eng	Intermediate	4.080	83354
7	6	2011	The Power of...	Charles Dugg	eng	Intermediate	4.030	155977
8	7	1994	Midnight in th...	John Banwart	eng	Intermediate	3.900	167987
9	8	2012	Hopeless	Colleen Hoover	eng	Intermediate	4.340	189938
10	9	1805	A Little Prince...	Francesc Hod...	eng	Intermediate	4.200	198872
11	10	2004	The Truth Ab...	Sarah Dessen	en-US	Intermediate	4.130	178415
12	11	1964	The horse an...	C.S. Lewis	eng	Intermediate	3.900	189671
13	12	2010	Last Sacrifice	Richest Mead	eng	Famous	4.420	206782
14	13	1935	Little House ...	Laura Ingalls...	eng	Intermediate	4.180	195424
15	14	1922	The Velvetes...	Margery Willi...	eng	Intermediate	4.290	173469

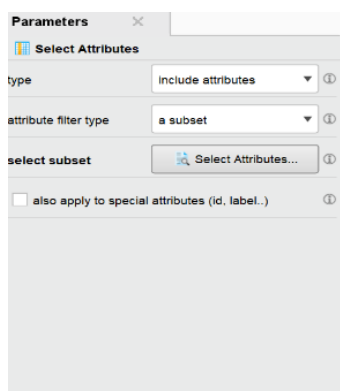
Gambar 4. Dataset yang sudah diimport

Yang dibagi menjadi data yang dibutuhkan, dari beberapa kolom dataset tersebut hanya 3 kolom dataset yang digunakan untuk mengetahui penjualan buku, kolom dataset itu berupa book name, genre, dan publisher. Dengan menseleksi dataset tersebut dengan menggunakan SELECT ATTRIBUTES.



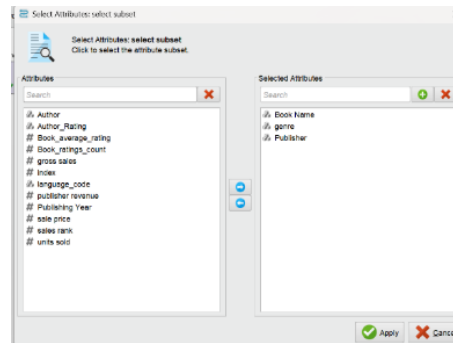
Gambar 5. Select Attributes

parameter yang digunakan select attribute untuk memisahkan dataset yang di pakai, dengan type : include attributes , attribute filter type : a subset dan tekan select subset: select attributes.



Gambar 6. Parameters

Dataset yang diseleksi hanya book name, genre, dan publisher, data ini menunjukan minat baca dan jumlah buku yang terjual bisa dilihat dari data yang tiga ini saja.



Gambar 7. Pembagian data yang digunakan

dari select attributes ini menghasilkan 1055 data yang bisa dipakai untuk melakukan algoritma FP-Growth.

Row No.	Book Name	genre	Publisher
1	Beowulf	genre fiction	HarperCollin...
2	Batman: Year...	genre fiction	HarperCollin...
3	Go Set a Wat...	genre fiction	Amazon Digit...
4	When You Ar...	fiction	Hachette Boo...
5	Daughter of ...	genre fiction	Penguin Gro...
6	Red Queen	genre fiction	Amazon Digit...
7	The Power of...	genre fiction	HarperCollin...
8	Midnight in th...	nonfiction	Hachette Boo...
9	Hopeless	genre fiction	HarperCollin...
10	A Little Prince...	genre fiction	Random Hou...
11	The Truth Ab...	genre fiction	Amazon Digit...
12	The horse an...	fiction	Penguin Gro...
13	Last Sacrifice	genre fiction	Amazon Digit...
14	Little House ...	genre fiction	HarperCollin...
15	The Velvetes...	genre fiction	Random Hou...
16	Zen and the ...	genre fiction	Amazon Digit...
17	The War of th...	genre fiction	Amazon Digit...

ExampleSet (1,055 examples,0 special attributes,3 regular attributes)

Gambar 8. Pembagian data yang digunakan

Setelah dilakukan pemilihan atribut dengan metode select attributes, proses dilanjutkan dengan pengujian menggunakan algoritma FP-Growth yang dikonfigurasi dengan format input berupa item yang dipisah di setiap kolom, menggunakan support sebagai syarat minimum dengan nilai minimum support sebesar 0.01, jumlah item minimum dalam satu itemset adalah satu, tanpa batas maksimum jumlah item dalam itemset, serta jumlah maksimum itemset yang dihasilkan mencapai satu juta.



Gambar 9. Algoritma FP-Growth Dengan hasil menunjukkan perbedaan nilai support dari berbagai pasangan item berukuran dua, di mana setiap kombinasi item 1 dan item 2 mencerminkan seberapa sering kedua item tersebut terjadi secara bersamaan dalam catatan transaksi, sehingga bisa digunakan guna mengidentifikasi pola pembelian konsumen, seperti kecenderungan genre tertentu muncul bersama penerbit tertentu atau layanan digital, yang pada akhirnya membantu pengambilan keputusan dalam strategi pemasaran dan rekomendasi produk.

Size	Support	Item 1	Item 2
1	0.768	genre fiction	
1	0.561	Amazon Digital Services, Inc.	
1	0.16	nonfiction	
1	0.113	Random House LLC	
1	0.1	Penguin Group (USA) LLC	
1	0.066	HarperCollins Publishers	
1	0.063	Hachette Book Group	
1	0.057	fiction	
1	0.051	Simon and Schuster Digital S...	
1	0.037	Macmillan	
1	0.014	children	
2	0.45	genre fiction	Amazon Digital
2	0.082	genre fiction	Random House LLC
2	0.083	genre fiction	Penguin Group (USA)
2	0.043	genre fiction	HarperCollins
2	0.045	genre fiction	Hachette Book Group
2	0.035	genre fiction	Simon and Schuster
2	0.029	genre fiction	Macmillan
2	0.076	Amazon Digital Services, Inc.	nonfiction
2	0.025	Amazon Digital Services, Inc.	fiction
2	0.01	Amazon Digital Services, Inc.	children
2	0.025	nonfiction	Random House LLC
2	0.017	nonfiction	HarperCollins
2	0.011	nonfiction	Hachette Book Group
2	0.011	nonfiction	Simon and Schuster

Gambar 10. Hasil Algoritma FP-Growth

Support tertinggi dalam tabel tersebut jatuh pada kolom "Item 1", yaitu pada item "genre fiction" dengan nilai support sebesar 0.768, yang berarti item ini paling sering muncul dalam data transaksi dibandingkan item-item lainnya, baik secara individu maupun dalam kombinasi dengan item lain di kolom "Item 2". Kolom "Item 2" pada

baris ini kosong, menandakan bahwa "genre fiction" muncul sendiri sebagai item yang dominan. Jadi, bisa disimpulkan bahwa item 1-lah yang memiliki support tertinggi. Dan ditambahkan lagi Operator Create Association Rules untuk menentukan tingkat kepercayaan (confidence) antar item.



Gambar 11. Algoritma FP-Growth yang ditambahkan create association rules

Jika proses ingin dibuat lebih optimal, berikut beberapa langkah perbaikan:

1. mengubah tipe data sudah diubah menjadi *transaksional* sebelum masuk ke FP-Growth.
2. Menyetel nilai *minimum support* dan *minimum Confidence* agar aturan yang dihasilkan tidak terlalu sedikit atau terlalu banyak.
3. Menambahkan operator *Result* atau *Table View* untuk menampilkan frequent itemset dan aturan secara lebih jelas.
4. Menghapus atribut yang tidak diperlukan sejak awal agar proses lebih efisien.

Operator *Create Association Rules* digunakan dengan konfigurasi yang menekankan tingkat kepercayaan (*confidence*) sebagai acuan utama dalam pembentukan aturan. Parameter *criterion* diset ke *confidence*, sehingga aturan yang dihasilkan diprioritaskan berdasarkan tingkat keyakinan hubungan antar item. Nilai *minimum confidence* ditetapkan sebesar 0.8, yang berarti hanya aturan dengan tingkat kepercayaan minimal 80% yang akan dihasilkan. Ini memastikan bahwa setiap aturan yang muncul memiliki kekuatan dan reliabilitas tinggi dalam merepresentasikan pola hubungan antar item.

Selain itu, parameter *gain theta* diatur sebesar 2.0 untuk meningkatkan

sensitivitas dalam memilih aturan yang memberikan informasi tambahan atau nilai pengetahuan yang signifikan. Parameter *laplace k* diset ke 1.0 sebagai koreksi statistik untuk mengatasi bias pada data dengan frekuensi rendah. Dengan konfigurasi ini, proses pembentukan asosiasi lebih terfokus pada kualitas aturan, memastikan bahwa setiap rule yang dihasilkan memiliki hubungan yang kuat, terpercaya, dan dapat digunakan secara valid dalam analisis serta pengambilan keputusan. dengan hasil menunjukkan dukungan (Support) antar item dan kepercayaan (confidence) item, pada table hasil association rule.

Tabel 1. hasil association rules

NO	Premises (jika)	Conciusion (Maka)	Support	Confidence
1	Penguin Group (USA) LLC	genre fiction	0.083	0.83

Maka Penguin Group (USA) LLC memiliki kecenderungan menerbitkan buku bergenre fiction, sehingga pola ini relevan untuk analisis pemasaran buku, rekomendasi kategori, dan segmentasi penerbit, serta dengan tingkat confidence yang tinggi (83%) dapat dijadikan indikasi kuat dalam penentuan strategi rekomendasi buku maupun prediksi genre berdasarkan penerbit.

**SIMPULAN (PENUTUP)**

Berdasarkan penerapan algoritma FP-Growth pada data penjualan *kitab*, penelitian ini membuktikan bahwa metode tersebut mampu mengidentifikasi pola pembelian pelanggan secara efektif sehingga dapat mendukung berbagai keputusan strategis bisnis. Seluruh proses yang dilaksanakan melalui tahapan KDD berhasil menghasilkan aturan asosiasi yang relevan, misalnya pola hubungan antara penerbit tertentu dengan jenis kitab yang mereka edarkan. Hasil analisis menunjukkan bahwa terdapat kategori kitab tertentu yang paling dominan dalam transaksi, dan penerbit tertentu memiliki

hubungan kuat dengan kategori tersebut, yang dibuktikan melalui nilai confidence yang tinggi. Temuan ini dapat dimanfaatkan untuk merancang strategi pemasaran, menentukan rekomendasi produk, menyusun paket promosi, serta melakukan segmentasi pasar secara lebih akurat. Selain itu, penggunaan RapidMiner membantu mempercepat proses analisis dan mempermudah visualisasi pola. Pada penelitian selanjutnya, variabel tambahan seperti harga kitab, rating pembaca, dan parameter penjualan lain dapat disertakan agar hasilnya lebih komprehensif. Penelitian lanjutan juga dapat menambahkan metode evaluasi lain guna meningkatkan akurasi. Dengan demikian, interpretasi pola asosiasi dapat diperoleh secara lebih tepat.

**DAFTAR PUSTAKA**

Jiawei Han, Micheline Kamber *Data Mining: Concepts and Techniques* karya, dan Jian Pei (edisi ke-3, 2012)

Sari, N. L., & Saputra, A. (2021). Analisis pola pembelian konsumen untuk optimasi inventaris menggunakan algoritma FP-Growth di UD Tritunggal Jaya. *Jurnal Informatika dan Teknologi (JITET)*, 10(2), 45–52.

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (edisi ke-3). Elsevier / Morgan Kaufmann

Suryadi, A., & Priyono, D. (2021). *Penerapan Association Rule dengan Algoritma FP-Growth untuk Transaksi Penjualan Produk pada Toko Aciro Berbasis Web*. Seminar Nasional Inovasi Teknologi dan Aplikasi (SENAFTI), Universitas Budi Luhur.

Retno Tri Vlandari. *Data Mining: Teori dan Aplikasi RapidMiner*. Gava Media, 2017. Laporan/liberatur di UTDI tentang penggunaan RapidMiner dalam data mining.

Prasetyo, F., & Hasugian, H. (2023). *Analisis pola pembelian produk makanan menggunakan algoritma FP-Growth untuk strategi penjualan*.

- Idealis: Indonesia Journal Information System*, 7(1), 11–20.
- Febrianty, N., Zaidiah, A., & Irmanda, H. N. (2024). Analisis minat konsumen pada produk CV Jakarta Powder Drink menggunakan metode asosiasi algoritma FP-Growth. *Journal of Information System (JOINS)*, 9(2), 94–103
- Subaegi, A., Faqih, A., & Anam, K. (2023). Implementasi data mining asosiasi pada data transaksi Es Teh Solo Jatibarang menggunakan algoritma FP-Growth dalam menentukan strategi penjualan. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(6), 3244–3251
- Sugiyono. (2019). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Rustam, C., Sari, R. P., & Hidayat, R. (2024). Analisis Pola Pembelian Konsumen Menggunakan Algoritma FP-Growth pada Data Penjualan. *Jurnal Teknologi dan Sistem Informasi*, 12(1), 45–52.
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann
- Sugiyono. *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta, 2017. Halaman 402.
- Han, Jiawei; Kamber, Micheline; Pei, Jian. *Data Mining: Concepts and Techniques*. 3rd Edition. San Francisco: Morgan Kaufmann, 2011. Halaman 19
- Han, Jiawei; Kamber, Micheline; Pei, Jian. *Data Mining: Concepts and Techniques*. 3rd Edition. San Francisco: Morgan Kaufmann, 2011. Halaman 83–84.
- Pang-Ning Tan; Michael Steinbach; Vipin Kumar. *Introduction to Data Mining*. Boston: Pearson Addison Wesley, 2006. Halaman 67.
- Fayyad, Usama; Piatetsky-Shapiro, Gregory; Smyth, Padhraic. *From Data Mining to Knowledge Discovery in Databases*. AI Magazine, 1996. Halaman 41.
- Hand, David; Mannila, Heikki; Smyth, Padhraic. *Principles of Data Mining*. Cambridge: MIT Press, 2001. Halaman 16.