

# PENERAPAN MARKET BASKET ANALYSIS (MBA) DATA MINING MENGGUNAKAN METODE ASOSIASI APPRIORI DAN FP-GROWTH UNTUK STRATEGI *BUNDLING* PADA WAN CAFFEINE ADDICT YOGYAKARTA

**Dony Satriyo Nugroho, Nur Islahudin, Vivi Normasari, Salsabiila Zaiima Al Hakiim**

Program Studi Teknik Industri,, Universitas Dian Nuswantoro, Jalan Nakula I No. 5-11, Pendrikan Kidul, Semarang, 50131

Email: [dony.satriyo.nugroho@dsn.dinus.ac.id](mailto:dony.satriyo.nugroho@dsn.dinus.ac.id)

**Abstrak--** Banyaknya *coffe shop* baru yang bermunculan setiap tahunnya, membuat para pelaku usaha *coffe shop* harus berinovasi dan menyediakan produk yang sesuai dengan preferensi konsumen, dimana salah satu preferensi konsumen adalah adanya paket produk serta diskon terhadap produk tertentu. Dengan penerapan teknologi informasi memiliki peran penting di dalamnya pemilik bisnis *coffe shop*, yaitu memudahkan pemilik bisnis dalam melakukan rekapitulasi dan pengolahan data yang akan membantu pemilik bisnis kafe mengambil keputusan yang berkaitan dengan peningkatan bisnisnya. salah satu metode yang digunakan untuk pengambilan keputusan adalah data mining Asosiasi. Asosiasi dapat digunakan untuk mencari keterkaitan antar produk sehingga mampu meningkatkan minat beli konsumen dengan cara *bundling* dan penetapan diskon terhadap pembelian produk-produk yang saling terkait. *Wan Caffeine Addict* yang merupakan sebuah *coffe shop* di Yogyakarta memiliki 684 transaksi dan 50 jenis produk terjual pada bulan desember 2022. Menggunakan metode asosiasi apriori dan FP-growth menggunakan parameter dengan tingkat support 3% dan confidence 10% . Dari hasil pengolahan diperoleh association rule produk Latte Hazelnut Large dan Wannabe memiliki confidence sebesar 0.7 dan lift ratio 3.84. selain itu dari association rule dapat disimpulkan bahwa produk wannabe dapat di *bundling* dengan produk lain seperti Balado Sticks, Sus Choco, dan Cassava Chips.

**Kata Kunci** : Data Mining, Asosiasi, Apriori, Fp-Growth, Coffe Shop, Wan Caffeine Addict

**Abstract--** The emergence of numerous new *coffe shops* every year requires *coffe shop* business owners to innovate and provide products that match consumer preferences, where one of the consumer preferences is the availability of product packages and discounts on certain products. The application of information technology plays an important role in helping *coffe shop* owners to easily recapitulate and process data that will assist them in making decisions related to business improvement. One method used for decision making is association data mining. Association can be used to find the correlation between products so as to increase consumer purchasing interest through *bundling* and setting discounts on related products. *Wan Caffeine Addict*, a *coffe shop* in Yogyakarta, had 684 transactions and sold 50 types of products in December 2022. By using the apriori and FP-growth association methods with a support level of 3% and a confidence level of 10%, the most related product was found to be Flavored Latte Hazelnut Large with the Wannabe product with a confidence level of 0.7 or 70% and a lift ratio of 3.84. From the processing results obtained the association rule of Latte Hazelnut Large and Wannabe products have a confidence of 0.7 and a lift ratio of 3.84. In addition, from the association rule it can be concluded that wannabe products can be *bundled* with other products such as Balado Sticks, Sus Choco, and Cassava Chips.

**Keywords** : Data Mining, Association, Apriori, Fp-Growth, Coffe shop, Wan Caffeine Addict

## 1. PENDAHULUAN

Di era digital saat ini, teknologi informasi memegang peran penting dalam segala aspek kehidupan, termasuk dalam aspek bisnis *coffe shop*. Bisnis *coffe shop* adalah salah satu bisnis yang sedang tumbuh dan berkembang di Yogyakarta, yang merupakan daerah yang padat dengan kaum pelajar dan pekerja. Persaingan yang ketat membuat para pemilik *coffe shop* harus mencari cara untuk meningkatkan penjualan dan mempertahankan pelanggan mereka. Banyaknya *coffe shop* baru yang bermunculan setiap tahunnya, membuat para pelaku usaha *coffe shop* harus berinovasi dan menciptakan produk yang unik untuk bisa bersaing. Konsumen pun menjadi lebih banyak pilihan dan memiliki standar yang lebih tinggi terhadap kualitas produk dan pelayanan yang diterima.

Menurut hasil penelitian (Kouahla et al. 2022) yang dilakukan pada tahun 2021, faktor yang paling penting bagi konsumen dalam memilih preferensi *coffe shop* adalah kualitas produk dan service, dimana diantara faktor-faktornya adalah paket produk yang menarik, dan adanya potongan harga terhadap produk yang ditawarkan. Oleh karena itu pemilik bisnis harus dapat menentukan jenis produk apa yang sebaiknya di-*bundling* (paket). Selain itu strategi *bundling* dapat meningkatkan minat beli konsumen (Wijaya dan Kinder, 2020) yang mana akan membantu pemilik *coffe shop* untuk bertahan ditengah munculnya banyak *coffe shop* baru. Namun dalam prakteknya strategi *bundling* tidak selalu diminati oleh konsumen. Salah satu aspek yang harus diperhatikan ketika menerapkan strategi *bundling* adalah jenis produk yang di-*bundling* memiliki kesamaan, kegunaan, wakt, distribusi, keuntungan, atau komplementer (Yan et al.2014). Dengan kata lain pemilik *coffe shop* harus mengetahui produk-produk yang saling berkorelasi.

Penambangan data (data mining) adalah proses menemukan pola dan hubungan dalam dataset besar dan menggunakannya untuk membuat keputusan yang terinformasi. Dalam beberapa tahun terakhir, penggunaan teknik penambangan data telah menjadi semakin populer dalam industri makanan dan minuman, termasuk *coffe shop*. Dengan menerapkan ketersediaan teknologi dan data bisnis yang melimpah penerapan data mining dalam bisnis telah banyak diterapkan oleh pelaku bisnis untuk mempertahankan dan mengembangkans bisnis (Ernawati 2018). Data Mining membantu *coffe shop* memahami preferensi dan perilaku pelanggan, mengoptimalkan operasi mereka, dan membuat keputusan yang lebih terinformasi (Rhomadhona, Aprianti, and Permadi 2021).

Implementasi Data di *coffe shop* dapat dicapai melalui penggunaan penambangan aturan asosiasi. Metode asosiasi Apriori yang merupakan teknik data mining yang populer dan efektif yang mengidentifikasi himpunan item sering dalam data transaksi dan menemukan hubungan antar satu atau lebih itemset (Safii and Trydillah 2019). Metode asosiasi ini bisa digunakan untuk memberikan rekomendasi kepada pelanggan atau untuk menyoroti kombinasi menu populer pada papan menu. Pemilik bisnis *coffe shop* juga bisa menggunakan data untuk menentukan item menu yang kurang populer dan mempertimbangkan untuk menghapus atau mengganti item tersebut. Dengan menganalisis kebiasaan pembelian pelanggan, pemilik bisnis dapat memperoleh wawasan tentang produk yang sering dibeli bersama, sehingga dapat digunakan untuk mengambil keputusan tentang penempatan produk (Anwar et al. 2020)atau desain menu, maupun tawaran paket promosi . Dengan kemampuan data mining tersebut, penerapan data mining dapat meningkatkan

keuntungan dari sebuah bisnis (Yulianton 2008) termasuk diantaranya bisnis coffe shop.

Beberapa penelitian sebelumnya telah melakukan penerapan asosiasi data mining dalam bisnis makanan dan minuman, seperti di restoran (Anwar et al. 2020), ataupun *e-commerce*. penelitian tentang penerapan metode asosiasi algoritma Apriori untuk analisis penjualan di *e-commerce* dengan Tujuan mempelajari pola belanja pelanggan dan memahami kebiasaan pembelian mereka. Hasil studi menunjukkan bahwa algoritma Apriori dapat membantu dalam mengidentifikasi pola pembelian pelanggan di *e-commerce*. (Lourenco and Varde 2020) . metode apriori juga efektif dalam menemukan aturan asosiasi dalam data transaksi toko ritel online dan untuk mencari hubungan transaksi dari sebuah toko ritel online (Lestari and Hafiz 2020) penerapan data mining asosiasi dengan menggunakan metode apriori untuk toko ritel konvensional juga berhasil mengekstraksi pola transaksional di toko ritel dan kecenderungan pembelian produk yang dibeli secara bersama-sama dalam sebuah transaksi, sehingga dapat digunakan untuk membuat keputusan pemasaran yang lebih efektif dan efisien (Ashari et al. 2022). Selain menggunakan apriori metode asosiasi pada data mining adalah metode FP-growth yang merupakan metode pengembangan dari metode apriori, dimana metode FP- growth tidak diperlukan generate candidate untuk memperoleh *frequent itemset*, sehingga membutuhkan waktu pemrosesan yang lebih cepat (Widiastuti and Sofi 2014).

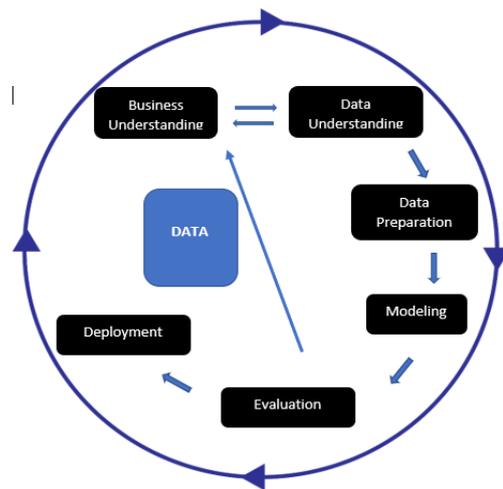
Dari berbagai penelitian tersebut menunjukkan bahwa penerapan asosiasi data mining dapat membantu meningkatkan penjualan dan memahami preferensi pelanggan. Namun, masih sedikit penelitian yang melakukan penerapan asosiasi data mining pada bisnis *coffe shop*. Oleh karena itu, penelitian ini bertujuan untuk mengisi

kekosongan tersebut dan memberikan referensi kepada pemilik kafe wan caffeine addict untuk memahami produk-produk yang saling berkaitan sebagai dasar pemilik wan caffeine addict untuk mengambil keputusan.

**2. TINJAUAN PUSTAKA**

**2.1 Data Mining**

Data mining merupakan sebuah teknik pembelajaran komputer (machine learning) yang berguna untuk menganalisis serta mencari sebuah pola dari data dan menjadikanya sebagai pengetahuan. Data mining juga dapat diartikan menjadi sebuah teknik untuk menemukan sebuah pola atau trend dari jumlah data yang sangat besar menggunakan metode statistik dan matematika. Data mining memiliki 5 peran, yaitu Estimasi, Forecast, Klasifikasi, Clustering, dan Asosiasi (Larose et al. 2012). Penerapan data mining sendiri memiliki berbagai langkah, diantaranya langkah yang paling populer adalah CRISP-DM (Cross Industry Standard Process for Data Mining):



Gambar 1. Alur Crisp-DM (Fadillah, 2015)

**2.2 Konsep Asosiasi**

Analisis asosiasi, merupakan salah satu peran dari Teknik pengolahan data pada

data mining. Analisis asosiasi dianggap sebagai fondasi untuk berbagai teknik data lainnya. Secara khusus, tahap tertentu dalam analisis asosiasi, yang disebut analisis pola frekuensi tinggi (frequent pattern mining), menarik perhatian banyak peneliti yang berupaya mengembangkan algoritma yang lebih efisien. (Wahyuni 2017). Salah satu algoritma yang efisien yaitu menggunakan metode Algoritma Apriori (Agrawal, Imieliński, and Swami 1993).

### 2.3 Algoritma apriori

Algoritma Apriori merupakan salah satu algoritma yang mencari frequent itemset dengan menggunakan teknik association rule. Algoritma ini menggunakan pengetahuan mengenai frekuensi atribut yang sudah diketahui sebelumnya untuk memproses informasi berikutnya (Wahyuni 2017). Pada Algoritma Apriori, kandidat yang mungkin muncul ditentukan dengan cara mempertimbangkan minimum *support* dan minimum *confidence*. Pada langkah awal, algoritma ini melakukan penggalan secara sistematis tanpa harus mengeksplorasi semua kandidat, sementara pada langkah berikutnya, ekstraksi dilakukan terhadap *rule* yang kuat. *Frequent itemset* mengacu pada himpunan item yang sering muncul bersama dalam suatu data transaksional. Sebagai contoh, jika item A dan B sering dibeli secara bersamaan di suatu toko retail. Setelah menemukan *frequent itemset*, kemudian algoritma ini melakukan penelitian lebih lanjut terhadap *knowledge* dari frequent itemset sebelumnya untuk menggali informasi lebih lanjut.. Algoritma Apriori menggunakan pendekatan iteratif dengan pencarian level-wise, di mana k-itemset digunakan untuk mencari (k+1)-itemset. Pada setiap iterasi, data i (kumpulan yang mengandung elemen i) yang sering muncul dihitung. Setiap iterasi terdiri dari dua langkah, yaitu *candidate generation* (pembentukan kandidat) dan *candidate counting and selection* (pemilihan

serta penghitungan kandidat). Parameter penting yang diperlukan untuk pembentukan rules dalam penerapan algoritma Apriori yaitu *Support* dan *Confidence*.

*Support* adalah nilai pengungjung atau persentase kombinasi sebuah item dalam database, Artinya adalah mencari kombinasi item yang memenuhi syarat dari nilai support minimum yang ada dalam basis data. Proses ini melibatkan pencarian nilai support menggunakan rumus yang dinyatakan sebagai berikut :

$$\text{Support (A)} = \frac{\sum \text{Transaksi Mengandung A}}{\text{Transaksi Total}}$$

(1)

(sumber : Maulidiya dan Jananto,2020)

Pencarian *support* terhadap kombinasi 2 item menggunakan formula berikut :

$$\text{Support (A, B)} = P(A \cap B)$$

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

(2)

(sumber : Maulidiya dan Jananto,2020)

Setelah semua pola frekuensi tinggi ditemukan, langkah berikutnya adalah mencari aturan asosiasi yang memenuhi syarat nilai *minimum confidence*. Ini dilakukan dengan menghitung nilai *confidence* untuk aturan asosiatif antara A dan B (Saefudin and DN 2019).

*Confidence* sering juga disebut sebagai akurasi dari suatu association rule dan mencerminkan sejauh mana hubungan yang kuat antar item dalam aturan asosiatif. Sebagai contoh, dalam aturan asosiasi  $A \Rightarrow B$ , *confidence* menunjukkan seberapa sering item B dibeli ketika konsumen membeli item A. Untuk menghitung nilai *confidence*, aturan tersebut diurutkan berdasarkan hasil *perkalian Support × Confidence*. Nilai *confidence* untuk aturan  $A \Rightarrow B$  dihitung menggunakan rumus berikut :

$$C = P(B | A)$$

$$Confidence(C) = \frac{\sum \text{Transaksi Mengandung A dan B}}{\sum \text{Transaksi Mengandung A}}$$

(3)

(sumber : Maulidiya dan Jananto,2020)

Nilai *expected confidence* dari aturan A B diperoleh dengan rumus berikut :

$$\text{Expected Confidence} = \frac{\sum \text{Transaksi Mengandung Support B}}{\sum \text{Transaksi}}$$

(4)

(sumber : Maulidiya dan Jananto,2020)

Untuk menentukan aturan asosiasi yang akan dipilih, langkahnya adalah mengurutkan aturan tersebut berdasarkan nilai *Support × Confidence*. Aturan dipilih sebanyak n aturan yang memiliki nilai ini terbesar. Prinsip kerja dasar dari algoritma ini adalah dengan mengembangkan *frequent itemset*, dimulai dari satu item, dan secara rekursif mengembangkan itemset tersebut menjadi dua item, tiga item, dan seterusnya, hingga *frequent itemset* tidak dapat dikembangkan lebih lanjut.

*Lift ratio* merupakan metode perhitungan yang lebih efektif untuk mengevaluasi kekuatan aturan asosiasi. Untuk menghitung nilai *Lift ratio*, langkahnya adalah membandingkan nilai *confidence* dibagi dengan *expected confidence*. Semakin menjauhi angka 1 maka aturan asosiasi dianggap semakin kuat. Berikut adalah rumus dari *Lift ratio* (Rizkiyanto 2020):

$$\text{Lift Ratio}(A, B) = \frac{\text{Support}(A \text{ dan } B)}{\text{Support}(A) * \text{Support}(B)}$$

(5)

### 2.3 Data

Data adalah informasi dalam bentuk symbol, angka, dan juga property yang diperoleh melalui sebuah pengamatan. Pada ilmu statistika data dibagi menjadi beberapa bentuk, yaitu:

- Ratio

Data yang diperoleh melalui pengukuran, di mana jarak antara dua titik pada skala sudah diketahui dan memiliki titik nol yang bersifat absolut contoh : Berat badan, Panjang, Tinggi badan, Umur, Jumlah uang.

- Interval

Data yang didapatkan melalui cara pengukuran, dimana data tidak mempunyai titik nol yang absolut dan jarak dua titik pada skala sudah diketahui.

Contoh : rentang pendapatan, Temperatur 10°C-1010°C, Umur 20-30 tahun

- Ordinal

Data yang didapatkan melalui pengelompokan kategori atau klasifikasi dan data tersebut bersigat urut.

Contoh :Tingkat kepuasan konsumen(sangat puas, puas, sedang, tidak puas

- Nominal

Data yang diperoleh dengan cara klasifikasi atau pengelompokan kategori, dan tidak ada atau tidak diketahui urutannya.

Contoh : spesies, Jenis kelamin, Jenis Penyakit, Nama kota

### 2.4 Transformasi Data

Dalam transformasi data, dilakukan perubahan tipe data atau konsolidasi tipe data ke dalam bentuk yang sesuai yang sesuai dengan tujuan atau metode data mining yang akan dipakai. Strategi untuk transformasi data meliputi hal-hal berikut ini:

- 1) Penghalusan / *Smoothing*, merupakan proses yang berfungsi untuk menghilangkan *noise* dari data. Beberapa Teknik yang dapat digunakan untuk melakukan *smoothing* meliputi regresi, pengelompokan, dan binning,

- 2) *Attribute Construction (feature construction)*, adalah proses menyusun atribut baru dan ditambahkan ke dalam set atribut untuk membantu proses *mining*.
- 3) Agregasi, merupakan operasi meringkas atau mengagregasi pada data. Sebagai contoh, data penjualan harian dapat diagregasi untuk menghitung total penjualan bulanan dan tahunan. Proses penghitungan jumlah total ini biasanya digunakan dalam pembangunan kubus data untuk analisis data pada tingkat abstraksi yang berbeda.
- 4) Normalisasi, melakukan penskalaan data dengan menggunakan metode sehingga berada dalam rentang yang lebih kecil, contoh data dengan rentang 1000 – 1000000 dapat dinormalisasi menjadi rentang -1 sampai 1.
- 5) Diskritisasi, proses merubah nilai mentah dari data bertipe numerik (misalnya, tinggi badan) menjadi tipe data interval (misalnya, 100-110, 11-50, dll.) dirubah ke dalam tipe ordinal (pendek, sedang, tinggi). Label-label tersebut, nantinya dapat disusun secara rekursif menjadi konsep-konsep yang lebih tinggi, membentuk sebuah hirarki konsep untuk tipe data numerik
- 6) Pembuatan hirarki konsep untuk data bertipe nominal melibatkan generalisasi atribut, seperti jalan, ke konsep yang lebih tinggi, seperti kota atau negara. Seringkali, banyak hierarki untuk atribut data bertipe nominal tersirat dalam skema database dan dapat otomatis didefinisikan pada tingkat definisi skema.

## 2.5 *One Hot-Encoding*

*One hot encoding* adalah salah satu metode untuk mengubah data untuk mempersiapkannya untuk sebuah algoritma dan mendapatkan prediksi yang lebih baik. Dengan *one-hot encoding* mengubah setiap nilai kategorikal menjadi kolom kategorikal baru dan memberikan nilai biner 1 atau 0 pada kolom-kolom tersebut. Setiap nilai bilangan bulat direpresentasikan sebagai vektor biner. Semua nilai adalah nol, dan indeks ditandai dengan angka 1. Data kategorikal mengacu pada variabel yang terdiri dari nilai label, misalnya, variabel "warna" dapat memiliki nilai "merah", "biru", dan "hijau". Pikirkan nilai seperti kategori yang berbeda yang terkadang memiliki urutan alami.

Beberapa algoritma pembelajaran mesin dapat bekerja secara langsung dengan data kategorikal tergantung pada implementasinya, seperti pohon keputusan, tetapi sebagian besar membutuhkan variabel input atau output berupa angka, atau nilai numerik. Ini berarti bahwa setiap data kategorikal harus dipetakan ke bilangan bulat.

Beberapa algoritma dapat bekerja dengan data kategorikal secara langsung; misalnya, *Decision Tree* dapat dipelajari secara langsung dari data kategorikal tanpa perlu transformasi data (hal ini bergantung pada implementasi spesifik). Sebaliknya, banyak algoritme pembelajaran mesin yang mengharuskan semua variabel input dan output berupa angka.

Dalam kasus ini, *one hot encoding* sangat membantu karena mengubah data kategorikal menjadi numerik; dengan kata lain: mengubah string menjadi angka sehingga kita dapat menerapkan algoritme Machine Learning tanpa masalah. Untuk variabel kategorikal yang tidak memiliki hubungan ordinal seperti itu, pengkodean bilangan bulat tidak cukup.

Dalam hal ini, *one hot encoding* dapat diterapkan pada representasi bilangan

bulat. Di sinilah variabel yang dikodekan dengan bilangan bulat dihapus dan variabel biner baru ditambahkan untuk setiap nilai bilangan bulat yang unik. Contoh *one-hot* encoding dapat dilihat pada Tabel 1. Dan Tabel 2.

Tabel.1 Contoh Transaksi

Nomor Transaksi	Produk terjual
1	kopi, gula, tepung
2	gula, teh
3	Teh, Tepung

Hasil encoding Transaksi

Tabel.2 Encoding Contoh Transaksi

No.	Kopi	Gula	Tepung	Gula	Teh
1	1	1	1	0	0
2	0	1	0	1	0
3	0	0	1	0	1

**2.6 Algoritma FP-Growth**

Algoritma FP-Growth adalah salah satu dari beberapa metode *pattern mining* yang memiliki dampak signifikan, di mana pola (seperti itemset, urutan, substruktur, subpohon, dan urutan) sering ditemukan. Metode pertumbuhan pola frekuensi ini mengekstraksi pola dari kumpulan data dengan pendekatan *divide-and-conquer*. Algoritma ini pertama-tama mengidentifikasi kumpulan pola frekuensi dengan ukuran 1, dan untuk setiap pola tersebut, melakukan proyeksi (atau kondisional) pada database dengan mengatur partisinya, dan kemudian secara rekursif mengekstraksi pola dari database yang diproyeksikan. Dengan mendekomposisi kumpulan data secara progresif menjadi subset-subset yang berkaitan dengan pola yang semakin kecil,

metode pertumbuhan pola efektif mengurangi ruang pencarian, menghasilkan efisiensi, dan mencapai skalabilitas yang tinggi.

Kecepatan FP-Growth melebihi algoritma Apriori karena prinsip dasarnya adalah pembangunan pohon dalam pencarian frequent itemsets. Dengan menggunakan struktur FP Tree, algoritma FP-Growth dapat secara langsung menggali *frequent itemset* dari *FP Tree* tersebut. *FP Tree*, yang merupakan singkatan dari *Frequent Pattern Tree*, merupakan suatu struktur pohon yang dibentuk dari kumpulan item awal dalam database. Tujuan dari *FP Tree* adalah untuk mengekstraksi pola yang paling sering muncul. Setiap node dalam pohon FP merepresentasikan satu item dari kumpulan item, dengan node root merepresentasikan null, dan node-node yang lebih rendah merepresentasikan itemsets. Hubungan antara node-node, yang menggambarkan itemset dengan itemset lainnya, dipertahankan ketika pohon dibangun. Langkah Metode FP-Growth adalah sebagai berikut :

- 1) Langkah awalnya adalah melakukan pemindaian database untuk mengidentifikasi kejadian itemsets dalam database, yang mirip dengan langkah pertama dalam algoritma Apriori. Perhitungan 1-itemset dalam database disebut sebagai support count atau frekuensi 1-itemset.
- 2) Langkah berikutnya adalah membuat FP Tree atau Frequent Pattern Tree, yang melibatkan pembuatan akar pohon dengan akar yang direpresentasikan oleh null
- 3) Langkah ketiga adalah melakukan pemindaian database dan memeriksa transaksi. Transaksi pertama diperiksa untuk menentukan kumpulan item yang terdapat di dalamnya. Itemset

dengan jumlah maksimum diambil dan ditempatkan di bagian atas. *Itemset* berikutnya dengan jumlah yang lebih rendah diambil, dan seterusnya. Dapat diamati bahwa cabang dari *FP Tree* atau *Frequent Pattern Tree* dibangun berdasarkan kumpulan item transaksi dengan urutan hitungan yang menurun.

- 4) Langkah berikutnya adalah memeriksa bagian transaksi dalam database. *Itemset* diurutkan dalam urutan hitungan yang menurun. Jika terdapat *itemset* dalam transaksi ini yang sudah ada di cabang lain (seperti pada transaksi pertama), maka cabang transaksi ini akan berbagi awalan yang sama hingga mencapai *root*, dan jumlah *itemset* akan bertambah, sebagaimana yang terjadi pada transaksi sebelumnya. Node umum akan dibentuk, dan jumlah node baru akan bertambah 1 saat node tersebut dibuat dan dihubungkan sesuai dengan transaksi.
- 5) Langkah berikutnya adalah melakukan penambangan pada *FP Tree* atau *Frequent Pattern Tree* yang telah dibuat. Pada langkah ini, node terendah dan link dari node terendah diperiksa terlebih dahulu. Node terendah merepresentasikan panjang pola frekuensi 1. Dari sini, jalur di *FP Tree* diikuti. Jalur ini disebut sebagai basis pola bersyarat. Dalam pembuatan *FP Tree* Bersyarat, yang terbentuk oleh jumlah *itemset* di jalur *Itemset* yang memenuhi ambang dukungan, dipertimbangkan dalam *FP Tree* Bersyarat. Pola Frekuensi dihasilkan dari *FP Tree* Bersyarat.

### 3. METODE PENELITIAN

#### Waktu dan Tempat Penelitian

Penelitian dilaksanakan di Wan Caffeine Addict yang beralamat di Jl,Pandean Sari No.6, Kentungan, Condongcatur, Kec. Depok, Kabupaten Sleman, Daerah Istimewa Yogyakarta dengan data yang digunakan adalah data transaksi penjualan Wan Caffeine Addict selama bulan Desember 2022.

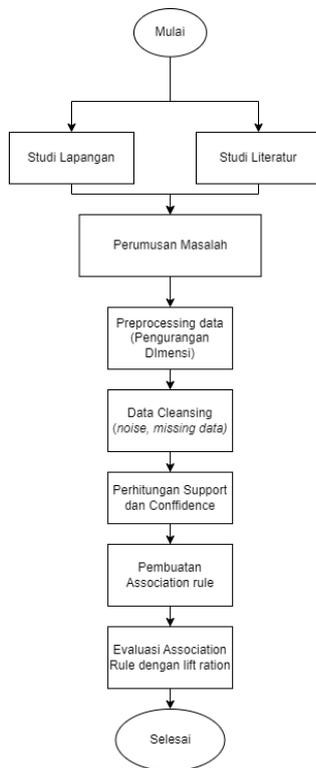
#### Pengambilan Data

Data yang digunakan merupakan data transaksi penjualan pada bulan Desember 2022 dengan total keseluruhan transaksi penjualan adalah 684 transaksi dengan jenis item yang dijual 50 jenis. Berikut adalah alur dalam pengolahan data

1. Melakukan identifikasi dan pemahaman terkait data yang dimiliki.
2. Tahap berikutnya adalah melakukan preprocessing data. Pada tahap ini dilakukan pemilihan dan pengurangan atribut data yang akan diolah menggunakan metode apriori dan fp-growth.
3. Data Cleansing yang digunakan untuk membersihkan data, terutama terkait data-data yang kosong atau terkait pesanan yang tidak tercatat.
4. Transformasi data menggunakan *one-hot encoding*, bertujuan untuk merubah data categorical menjadi numerik, dengan cara membuat seluruh jenis item yang dijual menjadi atribut dan memberikan nilai 0 atau 1 untuk setiap transaksi pada masing-masing atribut.
5. Perhitungan *Support* dilakukan untuk menyeleksi item apa saja yang akan diolah, pada kasus ini digunakan *support* sebesar 3% untuk kedua metode yang digunakan
6. Perhitungan *confidence* untuk mencari produk yang memiliki

- keterkaitan, digunakan nilai *confidence* minimal sebesar 10%
7. Pembentukan association rule, dari nilai *support* dan *confidence* akan diperoleh item mana saja yang memiliki hubungan dalam penjualan
  8. Evaluasi association rule bertujuan untuk melakukan apakah association rule yang terbentuk hanya merupakan sebuah kebetulan atau tidak.

Tahapan keseluruhan dalam penelitian ini dapat dilihat pada Gambar 2



Gambar 2. Diagram Alur Penelitian

#### 4 HASIL DAN PEMBAHASAN

Data yang digunakan pada penelitian ini adalah data transaksi pada bulan Desember 2022. pada bulan desember 2022 bulan tersebut terdapat 50 jenis produk yang dijual

Tabel 3. Item Produk Wan Caffeine Addict

Americano Large	Chiki Balls
Americano Reguler	Dawn Large
Aren Latte Large	Dawn Reguler
Aren Latte Reguler	Espresso
Balado Sticks	Espresso Single Shot
By Wan Large	Flafored Latte Caramel
By Wan Reguler	Flafored Latte Caramel Large
Cappucino Large	Flafored Latte Hazelnut
Cappucino Reguler	Flafored Latte Hazelnut Large
Cassava Chips	Flafored Latte Raspbery
Chocolate Hot	Flafored Latte Vanilla
Chocolate Large	Flafored Latte Vanilla Large
Chocolate Reguler	KaWAN Large
KaWAN Reguler	Red Velvet Large
Latte Large	Red Velvet Regular
Latte Reguler	Sunrise Large
Makaroni	Sunrise Regular
Matcha Hot	Sunset Large
Matcha Large	Sunset Regular
Matcha Regular	Sus Choco
MenaWAN Large	The Wan Large
MenaWAN Regular	The Wan Regular
Nightfall Large	Wannabe
Nightfall Regular	WANnabe Regular
nirWANA	
Red Velvet Hot	

Sedangkan untuk jumlah transaksi pada bulan Desember 2022 terdapat 684 transaksi baik yang dilakukan di tempat maupun secara online. pada tabel Tabel 4 merupakan contoh 10 transaksi yang terjadi

Tabel 4. 10 Transaksi awal bulan desember 2022

Waktu Transaksi	Nomor Transaksi	Jenis Produk yang dibeli
2022-12-01 11:01:06 +07:00	#9-2981	By Wan Regular
2022-12-01 11:27:10 +07:00	#9-2982	The Wan Large
2022-12-01 12:44:58 +07:00	#9-2983	The Wan Regular
		Wannabe
		Wannabe
2022-12-01 13:50:20 +07:00	#9-2984	By Wan Large
2022-12-01 13:58:32 +07:00	#9-2985	KaWan Regular
2022-12-01 15:28:32 +07:00	#9-2986	Wannabe
		Wannabe
2022-12-01 15:49:39 +07:00	#9-2987	Red Velvet Regular
		Red Velvet Regular
		Red Velvet Regular
		Sunrise Regular
	#9-2988	Americano Large

2022-12-01 16:24:11 +07:00		Americano Large
2022-12-01 16:26:13 +07:00	#9-2989	MenaWAN Regular
		Wannabe
		Matcha Regular
2022-12-01 16:49:29 +07:00	#9-2990	Matcha Regular
		Aren Latte Regular
		Aren Latte Regular

### 3.1 Encoding

Pada penerapan metode apripro untuk mencari pola keterkaitan antar produk harus dilakukan pengkodean dengan menggunakan metode one-hot encoding, contoh hasil pengkodean dari 10 transaksi diatas dapat dilihat pada tabel 5. Dimana pada tabel tersebut dapat dilihat bahwa seluruh jenis produk yang terjual menjadi variabel, dan 1 melambangkan produk tersebut terjual pada sebuah transaksi, dan 0 tidak terdapat penjualan pada sebuah transaksi. dari total hasil pengkodean diperoleh tabel matrik berukuran 684 (jumlah transaksi) x 50 (jenis Varian Produk).

Tabel 5. Pengkodean 10 transaksi awal bulan desember 2022

Americano Large	Aren Latte Regular	By Wan Large	By Wan Regular	KaWAN Regular	Matcha Regular	MenaWAN Regular	Red Velvet Regular	Sunrise Regular	The Wan Large	The Wan Regular	Wannabe
0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	1	1
0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	1	1	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0	0	0	1
0	1	0	0	0	1	0	0	0	0	0	0

### 3.2 Perhitungan Support

Dari 50 jenis produk tidak semua jenis produk dilakukan pengolahan asosiasi, melainkan hanya produk yang memiliki support diatas 3% dan menggunakan dynamic minimum support value dengan jumlah minimal itemsets 100, max number of retries 15 dan requirement decrease factor 0.9 pada metode fp-growth . Hal ini dikarenakan produk-produk yang ditawarkan oleh pihak kafe tidak semua jenis produk memiliki tingkat penjualan yang baik, sehingga ketika dilakukan analisis data tidak akan memberikan dampak yang cukup signifikan terhadap tingkat penjualan produk. Produk-produk yang memenuhi perhitungan nilai support pengolahan data dilakukan menggunakan bantuan software rapidminer

perhitungan support untuk produk wannabe pada 684 transaksi, dengan transaksi yang mengandung wannabe sebanyak 195 adalah sebagai berikut

$$Support\ W = \frac{\sum\ Transaksi\ mengandung\ W}{\sum\ Transaksi}$$

$$Support\ Wannabe = \frac{195}{684}$$

$$Support\ Wannabe = 18,13\%$$

### 3.3 Perhitungan Confidence

Perhitungan confidence digunakan untuk mengetahui keterkaitan antara satu produk dengan produk lain. Perhitungan keterkaitan antar dua produk, dimulai dari perhitungan confidence terhadap dua produk yang memiliki support lebih dari 3%. perhitungan confidence antara produk flavored hazelnut latte terhadap produk wannabe adalah sebagai berikut.

$$Confidence\ (W\ | \ FLHL) = \frac{Support(FLHL\ dan\ W)}{Support(FLHL)}$$

$$Confidence\ (W\ | \ FLHL) = \frac{1.9\%}{2.8\%}$$

$$Confidence\ (W\ | \ FLHL) = 0.7$$

data di-import ke dalam software rapidminer dan dilakukan preprocessing data berupa merubah tipe data menjadi tipe data dari tipe data integer ke dalam tipe data binomial untuk seluruh masing-masing variabel, dilanjutkan dengan operasi FP-growth dengan parameter support 0.03 (3%) dan dilanjutkan operasi Create Association Rule dengan parameter minimal confidence adalah 0.1 (10%) untuk membuat aturan asosiasi.

### 3.4 Perhitungan Lift Ratio

Lift ratio digunakan untuk melakukan validasi hasil dari association rule yang terbentuk. Lift ratio dapat menunjukkan apakah hasil dari association rule dari produk satu dengan produk lainnya merupakan sebuah kebetulan yang terjadi atau tidak. Perhitungan lift ratio untuk association rule Flavored Latte Hazelnut Large dengan produk Wannabe adalah

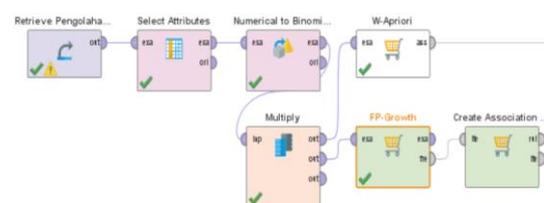
$$Lift\ Ratio\ (FLHL, W) = \frac{Support(FLHL\ dan\ W)}{Support(FLHL) * Support(W)}$$

$$Lift\ Ratio\ (FLHL, W) = \frac{0.019}{(0.027) * (0.183)}$$

$$Lift\ Ratio\ (FLHL, W) = 3.84$$

### 3.5 Pengolahan Menggunakan Rapid Miner

Pengolahan keseluruhan 684 transaksi menggunakan software rapidminer dengan menggunakan operator W-Apriori dan FP-Growth.



Gambar 3. Diagram proses asosiasi Apriori dan FP Growth pada rapidminers

Baik pada FP Growth dan W-Apriori menggunakan setup parameter min support 3% dan min confidence 10%

Gambar 4. Parameter minimal *support* 3%

Gambar 5. Parameter minimal *confidence* 10%

Dari hasil pengolahan menggunakan metode apriori dan FP-Growth dengan minimal *support* 3% dan minimal *confidence* 10% diperoleh aturan asosiasi sebagai berikut

### W-Apriori

```

Apriori
=====
Minimum support: 0.03 (23 instances)
Minimum metric <confidence>: 0.01
Number of cycles performed: 33

Generated sets of large itemsets:

Size of set of large itemsets L(1): 33
Size of set of large itemsets L(2): 2

Best rules found:

1. Flavored Latte Hazelnut Large=true 20 ==> Wannabe=true 14   conf:(0.7)
    
```

Gambar 6 Hasil Operasi Apriori

Sedangkan pengolahan menggunakan algoritma FP-Growth diperoleh *association rule* untuk 11

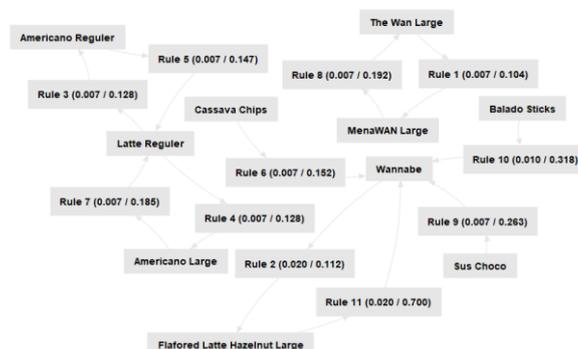
Tabel 6. hasil association rule FP-growth

Association Rules	Lift Ratio
[The Wan Large] --> [MenaWAN Large] ( <i>confidence</i> : 0.104)	2.74
[Wannabe] --> [Flavored Latte Hazelnut Large] ( <i>confidence</i> : 0.112)	3.84
[Latte Reguler] --> [Americano Reguler] ( <i>confidence</i> : 0.128)	2.58
[Latte Reguler] --> [Americano Large] ( <i>confidence</i> : 0.128)	3.25
[Americano Reguler] --> [Latte Reguler] ( <i>confidence</i> : 0.147)	2.58
[Cassava Chips] --> [Wannabe] ( <i>confidence</i> : 0.152)	0.83

[Americano Large] --> [Latte Reguler] ( <i>confidence</i> : 0.185)	3.25
[MenaWAN Large] --> [The Wan Large] ( <i>confidence</i> : 0.192)	2.74
[Sus Choco] --> [Wannabe] ( <i>confidence</i> : 0.263)	1.44
[Balado Sticks] --> [Wannabe] ( <i>confidence</i> : 0.318)	1.74
[Flavored Latte Hazelnut Large] --> [Wannabe] ( <i>confidence</i> : 0.700)	3.84

Baik dari metode apriori dan FP-growth diperoleh hasil *confidence* paling tinggi adalah Flavored Latte Hazelnut Large dengan produk wannabe yaitu sebesar 0.7 (70%). Akan tetapi metode FP growth mampu menampilkan *confidence* beserta item setnya.

Hasil dari diagram asosiasi terdapat pada gambar 7.



Gambar 7. Diagram asosiasi FP-Growth

Dari perhitungan hasil lifting rasio seluruh hasil aturan asosiasi memiliki lifting diantara kurang dari satu dan lebih dari satu. sehingga dapat disimpulkan hasil *confidence* pembelian sebuah produk dengan produk yang lainnya bukanlah sebuah kebetulan, dan hasil dari association rules yang diperoleh dapat dijadikan dasar pemilik kafe untuk mengambil keputusan.

## 4 KESIMPULAN

Dari hasil pengolahan data menggunakan metode FP-Growth dan apriori menggunakan minimal *support* 3% dan *confidence* 10% dengan total jenis penjualan produk 50 buah dan 684 transaksi, diperoleh

produk yang paling berkaitan adalah Wannabe dan Flafored Latte Hazelnut Large dengan tingkat *confidence* 70% serta lift ratio sebesar 3.84 yang menunjukkan bahwa produk tersebut sering dibeli oleh pelanggan. Disusul oleh produk Balado stick dan Wannabe.

Dari hasil analisis dilihat bahwa produk Wannabe dapat dibundling dengan produk seperti Flafored Latte Hazelnut, Balado Sticks, Sus Choco, dan Cassava Chips. Hal ini dikarenakan Association Rules yang diperoleh memiliki *confidence* cukup tinggi dan *lift ratio* yang bernilai tidak sama dengan satu

## 5 DISKUSI

Pada penelitian ini penulis hanya menggunakan data penjualan bulan desember dikarenakan keterbatasan rekapitulasi data, alangkah baiknya pada penelitian berikutnya menggunakan data satu siklus tahunan untuk hasil yang lebih baik. Serta kelemahan dari metode asosiasi FP growth ataupun apriori tidak menghiraukan jumlah pembelian produk yang saya pada tiap transaksi yang mengakibatkan *association rule* yang terbentuk tidak dapat merekomendasikan produk yang sama untuk dibundling. Hal ini membuka peluang untuk menggabungkan metode FP growth dan apriori dengan metode lainya agar didapatkan hasil yang lebih baik.

### Nomenclature

*FLHL* = Flafored Latte Hazelnut Large  
*W* = Wannabe

### Referensi

Agrawal, Rakesh, Tomasz Imieliński, and Arun Swami. 1993. "Mining Association Rules between Sets of Items in Large Databases." 207–16. doi: 10.1145/170035.170072.

Anwar, Muchamad Taufiq, Hindriyanto Dwi Purnomo, Mega Novita, and Clara Hetty Primasari. 2020. "Implementasi Metode Asosiasi Apriori Untuk Mengetahui Pola Beli Konsumen Dan Rekomendasi

Penempatan Produk Pada Swalayan Xyz." *Dinamik* 25(1):29–38. doi: 10.35315/dinamik.v25i1.7747.

Ashari, Imam Ahmad, Anggit Wirasto, Deny Nugroho Triwibowo, and Purwono Purwono. 2022. "Implementasi Market Basket Analysis Dengan Algoritma Apriori Untuk Analisis Pendapatan Usaha Retail." *MATRIK: Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer* 21(3):701–9. doi: 10.30812/matrik.v21i3.1439.

Ernawati, Iin. 2018. "Data Mining Sebagai Salah Satu Solusi Strategi Bisnis." *Informatik: Jurnal Ilmu Komputer* 14(1):9. doi: 10.52958/iftk.v14i1.367.

Fadillah, A. P. (2015). Penerapan Metode CRISP-DM untuk Prediksi Kelulusan Studi Mahasiswa Menempuh Mata Kuliah (Studi Kasus Universitas XYZ). *Jurnal Teknik Informatika Dan Sistem Informasi*, 1(3).

Kouahla, I., M. A. Yallese, S. Belhadi, K. Safi, and M. Nouioua. 2022. "Tool Vibration, Surface Roughness, Cutting Power, and Productivity Assessment Using RSM and GRA Approach during Machining of Inconel 718 with PVD-Coated Carbide Tool." *International Journal of Advanced Manufacturing Technology*. doi: 10.1007/s00170-022-09988-2.

Larose, Daniel T., D. Ph, Chantal D. Larose, and D. Ph. 2012. "Solutions to Chapter 1." *Algebra Through Practice* 29–50. doi: 10.1017/cbo9781139168274.007.

Lestari, Ade Fitria, and M. Hafiz. 2020. "Penerapan Algoritma Apriori Pada Data Penjualan Barbar Warehouse." *INOVTEK Polbeng - Seri Informatika* 5(1):96. doi: 10.35314/isi.v5i1.1317.

Lourenco, J., and A. S. Varde. 2020. "Item-Based Collaborative Filtering and Association Rules for a Baseline Recommender in E-Commerce." Pp. 4636–45 in *2020 IEEE International Conference on Big Data (Big Data)*.

- Maulidiya, H., & Jananto, A. (2020). Asosiasi Data Mining Menggunakan Algoritma Apriori Dan Fpgrowth Sebagai Dasar Pertimbangan Penentuan Paket Sembako.
- Rhomadhona, Herfia, Winda Aprianti, and Jaka Permadi. 2021. "Penerapan Data Mining Terhadap Data Penjualan Prioduk Kopi Menggunakan Algoritma Apriori." *Jurnal Sustainable: Jurnal Hasil Penelitian Dan Industri Terapan* 10(2):65–73. doi: 10.31629/sustainable.v10i2.3792.
- Rizkiyanto, Hendra. 2020. *Implementasi Algoritma Apriori Untuk Menganalisa Transaksi Pakaiian Di Toko Sahabat Busana Besuki*.
- Saefudin, Saefudin, and Septian DN. 2019. "Penerapan Data Mining Dengan Metode Algoritma Apriori Untuk Menentukan Pola Pembelian Ikan." *JSiI (Jurnal Sistem Informasi)* 6(2):36. doi: 10.30656/jsii.v6i2.1587.
- Safii, M., and Alrizca Trydillah. 2019. "METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi IMPLEMENTASI DATA MINING DALAM MENENTUKAN POLA PEMBELIAN OBAT DENGAN METODE ALGORITMA APRIORI." *METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi* 3(1):66.
- Wahyuni, Siti Tri. 2017. "Implementasi Data Mining Algoritma Apriori Pada Sistem Penjualan Roti Di Difa Rien's Bakery." 53(9):1689–99.
- Widiastuti, Dwi, and Nelly Sofi. 2014. "Analisis Perbandingan Algoritma Apriori Dan Fp-Growth Pada Transaksi Koperasi." *UG Jurnal Vol.* 8(01):21–24.
- Wijaya, A., & Kinder, L. (2020). Pengaruh Price Bundling dan Product Bundling terhadap Niat Membeli yang Dimoderasi oleh Barang Komplementaritas. *Jurnal Manajemen*, 17(1), 28-38.
- Yan, R., Myers, C., Wang, J., & Ghose, S. (2014). Bundling products to success: The influence of complementarity and advertising. *Journal of Retailing and Consumer Services*. <https://doi.org/10.1016/j.jretconser.2013.07.007>
- Yulianton, Heribertus. 2008. "Data Mining Untuk Dunia Bisnis Keputusan Informasi." *Jurnal Teknologi Informasi DINAMIK* XIII(1):9–15.