**MENGIMPLEMENTASIKAN ALGORITMA APRIORI DALAM SUPERMARKET SALES DATA**

1st Jovinka Aphellia Salva  
*Program Studi Ilmu Komputer  
Universitas Sumatera Utara*Medan, Indonesia   
jovinkav@gmail.com2nd Najwa Afifi Situmorang  
*Program Studi Ilmu Komputer  
Universitas Sumatera Utara*Medan, Indonesia  
najwaafifi121@gmail.com

3rd Dewi Sartika Br Ginting S.Kom., M.Kom  
*Program Studi Ilmu Komputer  
Universitas Sumatera Utara*Medan, Indonesia   
dewidintingdg90@gmail.com

***Abstrak***— Dengan menggunakan teknik penambangan aturan asosiasi, kami bertujuan untuk mendistribusikan kelompok item yang umum muncul dan menghasilkan aturan asosiasi yang bermakna yang dapat memberikan wawasan berharga tentang perilaku pembelian pelanggan.

Kumpulan data terdiri dari detail selanjutnya seperti ID faktur, lini produk, dan model pembayaran dan telah diproses sebelumnya untuk menyederhanakan proses utama. Analisis kami mengungkapkan bahwasanya nilai support ke Fashion accessories (0,178) yang merupakan nilai support tertinggi di antara lainnya.

Kata Kunci—asosiasi, apriori, penjualanan, data

# **PENDAHULUAN**

Algoritma apriori adalah proses penggalian informasi dari database dan seringkali menghasilkan elemen atau kumpulan elemen dan kandidat untuk membentuk asosiasi rule mining untuk memperoleh nilai dukungan minimum dan keyakinan minimum.

Untuk database yang cukup besar, algoritma apriori akan menghasilkan pola item/itemset yang sering digunakan dalam jumlah besar, karena algoritma ini perlu membuat kandidat dan mencatat database yang berulang.

Pada penelitian ini, kami menerapkan algoritma Aprior pada dataset toko supermarket yang berisi data penjualan transaksi supermarket. Tujuan utama penelitian ini adalah untuk menemukan kelompok produk umum yang sering dibeli konsumen secara bersamaan.Dengan mengidentifikasi perilaku pembelian ini, pemilik supermarket dapat memperoleh wawasan tentang taktik pemasaran dan teknik penjualan mereka untuk meningkatkan kinerja secara keseluruhan.

Sebelum menerapkan algoritma Aprior, kami melakukan penambangan data awal untuk memahami karakteristik kumpulan data dan mengidentifikasi informasi penting seperti produk terlaris, metode pembayaran terpopuler, dan penjualan berdasarkan jenis kelamin dan bulan. Tujuan dari langkah ini adalah untuk mendapatkan wawasan tentang pola pembelian pelanggan dan menyiapkan data untuk analisis lebih lanjut.

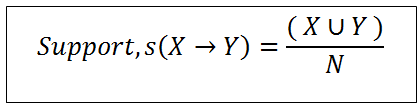
# **TEORI**

Apriori adalah algoritme untuk menambang kumpulan target dan mempelajari aturan asosiasi secara berulang melalui database relasional. Prosesnya dengan mengidentifikasi objek tertentu yang ditemukan berulang kali dalam database dan kemudian secara bertahap meningkatkan jumlah objek tersebut hingga tersedia untuk dianalisis**.**

Menambang produk yang sering digunakan dan aturan asosiasinya dicapai melalui penggunaan algoritma Apriori. Secara umum, algoritma apriori bekerja pada database yang berisi transaksi dalam jumlah besar. Misalnya barang konsumsi, tapi di Grand Bazaar.

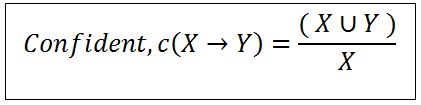
Support, confidence, dan lift menjadi tiga komponen utama dalam proses data mining asosiasi menggunakan algoritma apriori.

Bentuk persentase jumlah kemunculan untuk kombinasi item tertentu disebut dengan nilai Support(s).



Pentingnya nilai-nilai dukungan dalam aturan asosiasi ditekankan, karena tingkat dukungan yang rendah menunjukkan bahwa asosiasi jarang terjadi dalam kumpulan data (semua data peristiwa).

Menghitung persentase dari keakurasian aturan asosiasi yang akan dihasilkan disebut dengan nilai Confident(c).



Besarnya Y didefinisikan sebagai keyakinan yang tinggi untuk peristiwa yang mengandung X.

***Lift Ratio*** adalah parameter yang mengukur kekuatan aturan asosiasi yang diciptakan oleh nilai-nilai dukungan dan keyakinan.

# **METODOLOGI PENELITIAN**

Penelitian ini menerapkan algoritma apriori dari Data Mining untuk mengekstrak data terkait penjualan supermarket.

Data yang digunakan adalah catatan penjualan supermarket yang berisi informasi tentang seluruh transaksi penjualan, seperti produk yang dibeli, metode pembayaran dan demografi pelanggan.

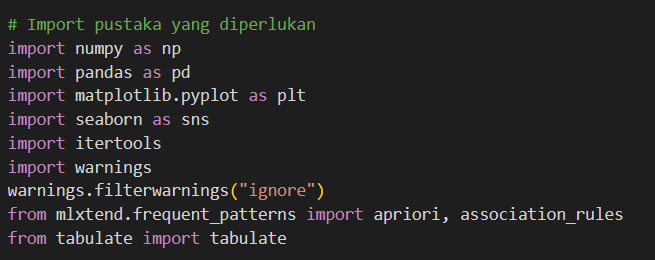
## Data dan Preprocessing

* 1. Sumber Data :

Dataset yang digunakan berasal dari file CSV Kaggle berjudul “supermarket\_sales.csv”. Data yang ada pada dataset ini mencakup berbagai atribut seperti ID transaksi, kota, anggota, jenis kelamin, produk yang dibeli, harga, jumlah, tanggal, metode pembayaran, dan lainnya.

* 1. Pembacaan Data :

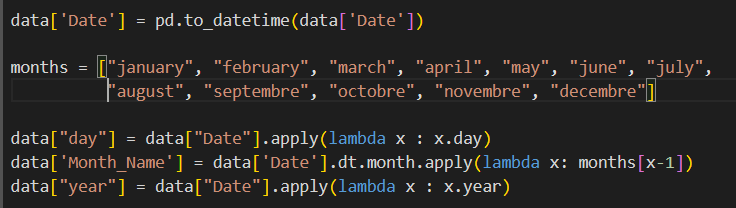
Dataset dibaca menggunakan pustaka `pandas`, `numpy`, `matplotlib.pyplot`, `seaborn`, `itertools` dan `warnings`dengan perintah berikut :



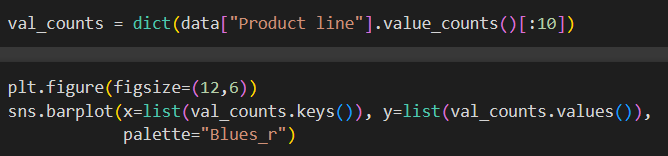
* 1. Preprocessing Data :

Data diproses untuk mempersiapkannya bagi analisis lebih lanjut.

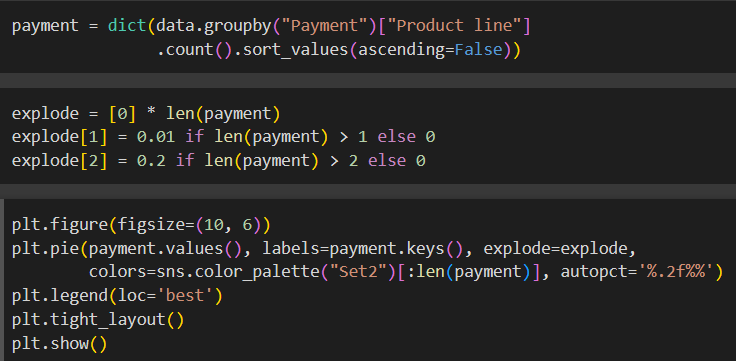
* Mengonversi kolom tanggal ke tipe datetime.
* Menambah kolom yang memisahkan informasi hari, bulan, dan tahun dari kolom tanggal.
* Menghapus kolom yang tidak diperlukan untuk analisis.



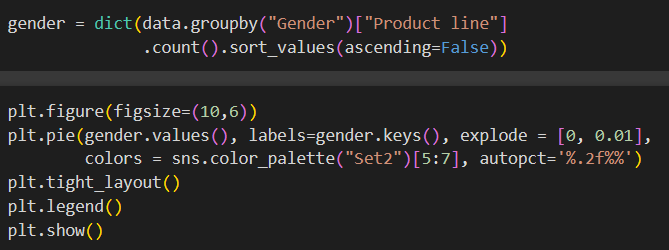
1. *Analisis Data Eksploratif* 
   1. Visualisasi Jumlah Produk Yang Dibeli



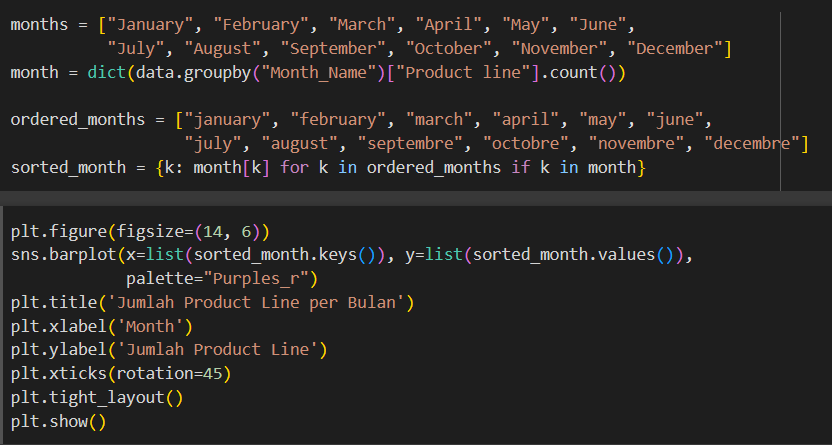
* 1. Visualisasi Fitur Metode Pembayaran



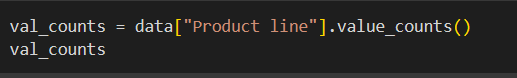
* 1. Visualisasi Fitur Jenis Kelamin



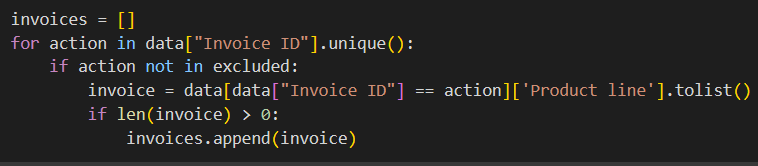
* 1. Visualisasi Fitur Bulan



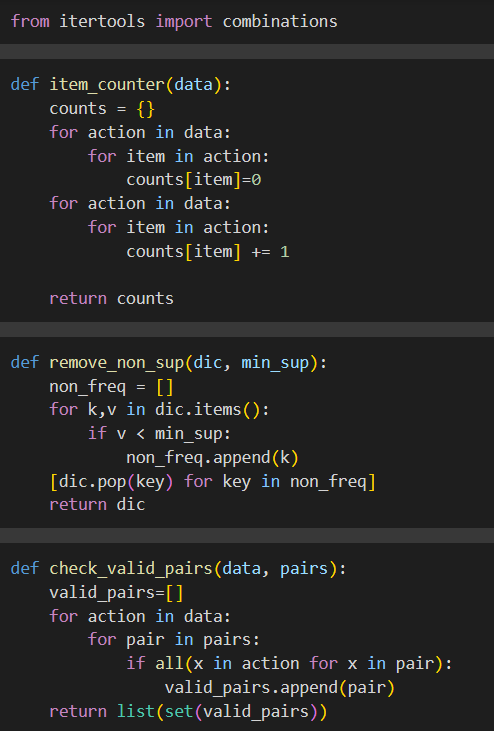
1. *Preprocess Data (untuk Penambangan Data Pola)*
   1. Penjumlahan Tiap Produk Yang Dibeli

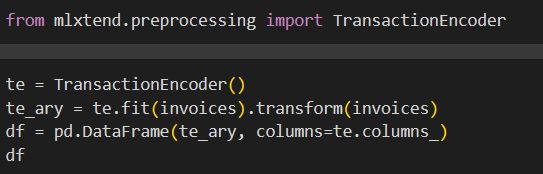


* 1. Penampilan ID Transaksi



1. *Implementasi Algoritma Apriori*



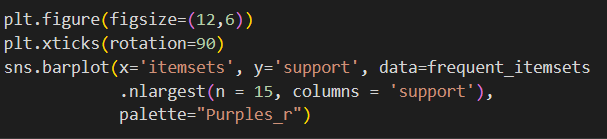
1. *Mengambil Hasil Akhir dan Membandingkannya Dengan Algoritma Yang Asli/Fakta*
   1. Menampilkan TransaksiEncorder
   2. Pembentukan Frequent Itemsets

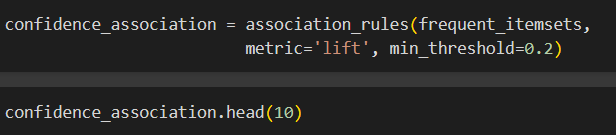
Frequent itemsets dibentuk dengan menggunakan fungsi `apriori` dari pustaka `mlxtend.frequent\_patterns` dengan minimal support yang telah ditentukan.

* 1. Pembentukan Nilai Terbesar pada Frequent Itemsets



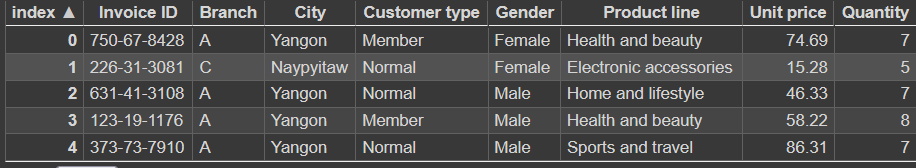
* 1. Visualisasi Gambar Frequent Itemsets



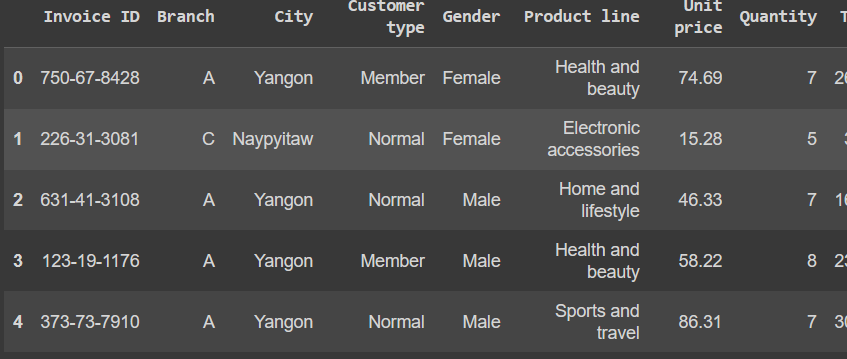
1. Interprestasi Aturan Asosiasu

## Units

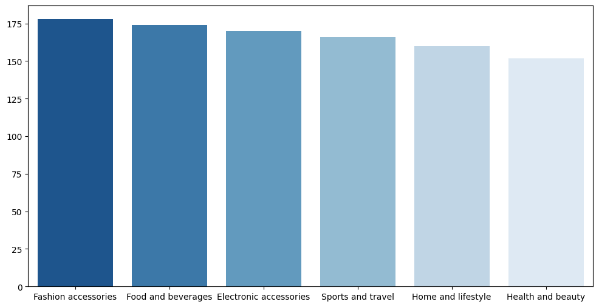
# **HASIL DAN PEMBAHASAN**



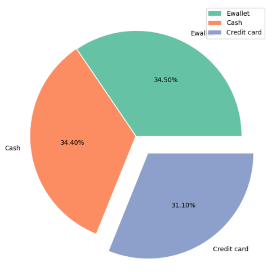
Gambar 1 : Data Belajar sudah Diambil dan Dijalankan



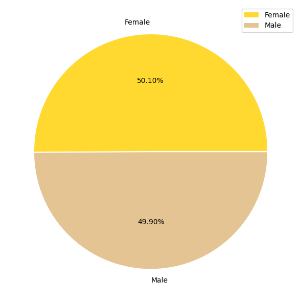
Gambar 2 : Preprocessing Data



Gambar 3 : Visualisasi Jumlah Produk Yang Dibeli

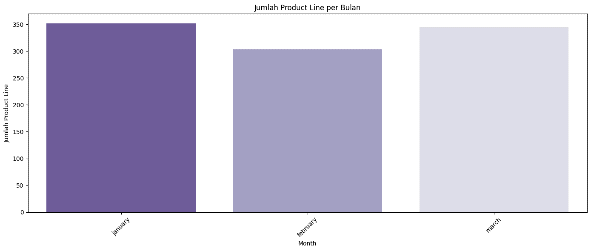


Gambar 4 : Visualisasi Fitue Metode Pembayaran

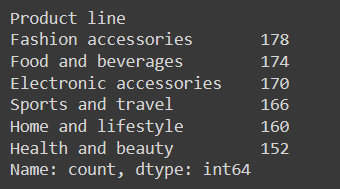


\\

Gambar 5 : Visualisasi Fitue Jenis Kelamin



Gambar 6 : Visualisasi Fitur Bulan



Gambar 7 : Jumlah Tiap Produk yang Dibeli



Gambar 8 : Tabel Transaksi Ercorder

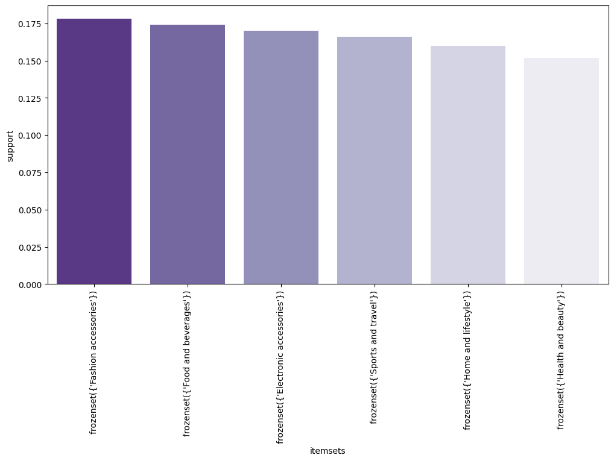


Gambar 9 : Tabel Support 1 Itemset



GGGGGGGG

Gambar 10 : Tabel Support



Gambar 11 : Gambar Figur dari Support

# **KESIMPULAN**

Kami menggunakan algoritma Aprior untuk menerapkan analisis penambangan pola berulang pada data supermarket penjualan acara dalam penelitian ini. Fungsi sepertmemasak"penghitung\_item", "*remove\_non\_sup*","*check\_valid\_pairs*","*pair\_counter*","*unique\_elements*", dan "apriori" dibuat untuk mengimplementasikan algoritme Apriori secara manual.

Analisis ini menunjukkan kelompok umum atau kelompok produk yang sering dibeli oleh pelanggan supermarket dalam jumlah besar dalam periode waktu yang berbeda (misalnya 2, 3, 4, 5 item kehabisan stok). Kumpulan item dengan panjang item 2 dan 3 mewakili pola pembelian yang paling sering diamati dalam data transaksi.

Penelitian ini tidak hanya mencari jumlah item yang sering dilihat, tetapi juga menghitung nilai kepercayaan antar item atau kelompok produk tertentu. Kepercayaan dihitung menggunakan fungsi equity\_on\_items dan equity\_on\_sets. Ini menunjukkan seberapa sering suatu produk atau lini produk dibeli ketika produk atau lini produk lain dibeli dalam transaksi yang sama.

Hasil penambangan model dan analisis kepercayaan berkelanjutan membantu supermarket memahami perilaku pembelian konsumen. Informasi ini dapat digunakan dalam strategi pemasaran, kampanye, penempatan produk atau pengembangan produk baru untuk memenuhi preferensi pelanggan. Secara keseluruhan, penelitian ini menunjukkan bahwa algoritma Aprior dapat menemukan pola belanja yang berguna untuk pengambilan keputusan bisnis melalui analisis pattern mining pada data transaksi supermarket.

# **DAFTAR PUSTAKA**

[**https://en-m-wikipedia-org.translate.goog/wiki/Apriori\_algorithm?\_x\_tr\_sl=en&\_x\_tr\_tl=id&\_x\_tr\_hl=id&\_x\_tr\_pto=tc**](https://en-m-wikipedia-org.translate.goog/wiki/Apriori_algorithm?_x_tr_sl=en&_x_tr_tl=id&_x_tr_hl=id&_x_tr_pto=tc)

[**https://www-javatpoint-com.translate.goog/apriori-algorithm?\_x\_tr\_sl=en&\_x\_tr\_tl=id&\_x\_tr\_hl=id&\_x\_tr\_pto=tc**](https://www-javatpoint-com.translate.goog/apriori-algorithm?_x_tr_sl=en&_x_tr_tl=id&_x_tr_hl=id&_x_tr_pto=tc)

[**https://colab.research.google.com/drive/13bAvG56tL1kY113duD9pCNr1s1qXFFUB?authuser=0#scrollTo=tnSj\_c8nI311**](https://colab.research.google.com/drive/13bAvG56tL1kY113duD9pCNr1s1qXFFUB?authuser=0#scrollTo=tnSj_c8nI311)