

## Implementasi *Mechine Learning Intelligence* Sistem Menggunakan Algoritma Apriori

<sup>1</sup>Cendra Wadisman, <sup>2</sup>Ahmad Junaidi, <sup>3</sup>Sintia

<sup>1,2</sup>Universitas Putra Indonesia “YPTK” Padang

<sup>3</sup>Universitas Alifah Padang

Email: [1cendra\\_wadisman@upiypk.ac.id](mailto:cendra_wadisman@upiypk.ac.id), [2ahmad\\_junaidi@upiypk.ac.id](mailto:ahmad_junaidi@upiypk.ac.id),  
[3sintia121994@gmail.com](mailto:sintia121994@gmail.com)

Corresponding Author : [cendra\\_wadisman@upiypk.ac.id](mailto:cendra_wadisman@upiypk.ac.id)

### **Abstract**

*This study aims to implement a data mining application to produce association rules between items in a transaction to purchase goods of various types of goods simultaneously by determining support of 40% and confidence of 50%. Thus, if there is a consumer buying a type of goods. The problem that arises in the K92-Mart minimarket is that there are often no sales of goods that consumers want or are out of stock because they do not pay attention to stock. To overcome this problem, a data mining application was created using the Apriori Algorithm. By implementing this apriori algorithm, it is hoped that it can produce a decision from sales data that aims to determine the pattern of purchasing goods that are often purchased by consumers and are strategic in sales.*

**Keywords:** Data Mining, Algoritma Apriori, Sale.

### **Pendahuluan**

Perkembangan Teknologi Informasi (TI) telah mengalami kemajuan yang sangat pesat, termasuk penggunaannya di sektor pemasaran. Komputerisasi telah diterapkan oleh banyak perusahaan, namun umumnya digunakan sebatas membuat suatu sistem penyimpanan data yang hanya digunakan untuk pencatatan transaksi dan mengolah data transaksi tersebut untuk pembuatan laporan pembelian atau penjualan. Bagi perusahaan yang hanya menjual beberapa jenis produk atau toko-toko kecil, laporan seperti itu mungkin sudah cukup bagi pemilik untuk melakukan analisis pada pasar dan mengambil keputusan. Produk-produk dalam skala besar, terdiri dari ratusan bahkan ribuan jenis produk, cara pengolahan data lebih sulit dilakukan (Dzulkarnaen, 2020).

Semakin berkembangnya persaingan dalam dunia bisnis khususnya dalam industri supermarket, minimarket dan swalayan menuntut para pengembang untuk menemukan suatu pola yang dapat meningkatkan penjualan dan pemasaran barang di swalayan, salah satunya adalah dengan pemanfaatan data transaksi. Ketersediaan data yang melimpah, kebutuhan akan informasi sebagai pendukung pengambilan keputusan untuk membuat solusi bisnis, dan dukungan infrastruktur di bidang teknologi informasi merupakan alasan dari lahirnya teknologi data mining. Masalah yang terjadi di Minimarket K92 yaitu manager kurang dalam peninjauan produk yang dijual, produk-produk apa saja yang dibutuhkan konsumen dan penyimpanan data-data kurang efektif (Nofitri & Eska, 2018).

## Landasan Teori

### **Hybrid Intelligence System (HIS)**

*Hybrid Intelligence System* (HIS) merupakan pendekatan yang menggabungkan kemampuan manusia dengan kecerdasan buatan (Artificial Intelligence) untuk menciptakan sistem yang lebih adaptif dan efektif dalam menyelesaikan masalah kompleks. Salah satu tantangan dalam HIS adalah bagaimana memanfaatkan data secara optimal untuk menghasilkan prediksi atau keputusan yang akurat. Dalam konteks ini, machine learning (ML) berperan penting sebagai salah satu komponen utama HIS (Araf Aliwijaya & Hanny Chairany Suyono, 2023).

### **Data Mining**

Data mining sendiri sering disebut sebagai *knowledge discovery in database (KDD)* adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola hubungan dalam set data berukuran besar. Output dari data mining ini dapat digunakan untuk pengambilan keputusan di masa depan. Salah satu teknik yang dikenal dalam data mining yaitu clustering. Berupa proses pengelompokan sejumlah data atau objek ke dalam *cluster* (group) sehingga setiap dalam *cluster* tersebut akan berisi data yang semirip mungkin dan berbeda dengan objek dalam *cluster* yang lainnya (Royal, 2024).

### **Algoritma Apriori**

Algoritma Apriori adalah algoritma untuk menemukan pola frekuensi tinggi. Algoritma apriori mempunyai kelebihan lebih sederhana dan dapat menangani data yang besar dalam penggunaan memori saat memproses data yang besar tentunya berpengaruh terhadap banyaknya item yang diproses serta mudah dipahami struktur kerja dan implementasinya (Katarina Sianturi et al., 2023).

Penelitian terdahulu dilaksanakan untuk klasifikasi kualitas air layak minum. Data yang digunakan diperoleh dari data water quality di kaggle. Data yang didapatkan mula-mula akan dilakukan proses data cleaning. Data cleaning dalam Algoritma apriori yang bertujuan untuk menemukan *frequent item sets* pada sekumpulan data. Algoritma apriori didefinisikan suatu proses untuk menemukan suatu aturan apriori yang memenuhi syarat minimum untuk *support* dan syarat minimum untuk *confidence* (Febriyanti & Zakaria, 2023).

Penelitian selanjutnya dilakukan oleh Analisis data dilakukan dengan cara membuat ringkasan dari sumber sumber yang didapatkan. Kesimpulan diambil dari hasil pengamatan dan analisis antara literatur dan fakta lapangan yang memperkuat data yang ada. Penjabaran hasil penelitian dilakukan secara deskriptif dan didukung dengan gambar-gambar untuk memperjelas hasil analisis (Wismoaji, 2024).

## Metode Penelitian

### **Algoritma Apriori**

Apriori adalah suatu algoritma yang sudah sangat dikenal dalam melakukan pencarian *frequent itemset* untuk mendapatkan *association rules*. Sesuai dengan namanya algoritma ini menggunakan *knowledge* mengenai *frequent itemset* yang telah diketahui sebelumnya, untuk memproses informasi selanjutnya. Apriori menggunakan pendekatan

secara intertite yang diketahui juga dengan level-wise dimana K-itemset digunakan untuk mencari (K+1) itemset. Pertamaakan dicari set *frequent* 1-itemset, set ini dinotasikan sebagai L1. L1 digunakan untuk menentukan L2, begitu seterusnya sampai *frequent itemset* tidak ada lagi (Supardi et al., 2023).

Dengan rumus sebagai berikut:

1. Analisis pola frekuensi tinggi

Nilai Support (nilai penunjang) merupakan presentase dari record-record yang mengandung kombinasi.

$$Support (A) = \sum \frac{Transaksi Mengandung (A)}{Total Transaksi (A)} \times 100$$

Penjelasan.

Support A = Menghitung Nilai Item Dalam Sebuah Dataset

Transaksi = Menghitung Jumlah Transaksi Yang Ada Pada Dataset

Total Transaksi = Jumlah Transaksi Dalam Dataset

2. Pembentukan Aturan Asosiasi

Akurasi dari suatu association rule sering disebut dengan confidence. Confidence (nilai kepastian) adalah kuatnya hubungan antara item dalam aturan asosiatif.

$$Confidence (A \Rightarrow B) = \frac{Support (A dan B)}{Support (A)} \times 100$$

Penjelasan.

Confidence A-B = Pengabungan Jumlah Datase

Support (A-B) = Gabungan Jumlah Transaksi Mengandung A dan B

Support A = Jumlah Transaksi Yang Mengandung Nilai A.

### Hasil dan Pembahasan

Dalam tahapan penelitian sistem dalam mengidentifikasi analisa data penjualan barang pada peneliti ini melakukan analisa data pada penjualan barang yang telah ada pada minimarket K92-Mart untuk mencari analisa frekuensi tinggi kombinasi item *support* dalam menentukan pola hubungan antara item-item pembelian.:

1. Identifikasi Data Barang

Pada tahapan ini adalah identifikasi data penjualan barang pada minimarket K92-Mart dapat dilihat pada tabel berikut:

**Tabel 1. Identifikasi Data Barang**

No.	Kode Barang	Nama Barang
1	BRG0001	Dunhill 16 Mild Putih
2	BRG0002	Aqua 600 ml
3	BRG0003	Capucino
4	BRG0004	Surya 12
5	BRG0005	Aqua 1500 ml
6	BRG0006	Sari Roti Krim Coklat
7	BRG0007	Sampoerna 16 Mild

Tahapan penelitian yang dijelaskan pada tabel 1 mencakup langkah-langkah untuk menyelesaikan masalah dalam Analisa data per transaksi pada tahap ini maksudnya adalah memisahkan transaksi ke-1 yaitu: dunhill 16 mild putih, sari roti, walls, minyak goreng, sari roti, sampurna. Untuk menentukan data transaksi penjualan pada suatu faktur dalam pembelian.

2. Analisa Data Frequence I-itemset

Pada tahapan ini adalah Analisa kombinasi frequent itemset pada tahap ini dilakukan analisa dilakukan dengan menggunakan kombinasi *I-itemset* dari data transaksi tersebut menghitung jumlah *support*-nya dengan min support 25% dan min confidence 50%.

**Tabel 2. Data Frequence I-Itemset**

No.	Atribut Barang 1	Jumlah	Support	keterangan
1	Dunhill 16 Mild Putih	1	25 %	Lolos
2	Aqua 600 ml	1	25 %	Lolos
3	Capucino	1	25 %	Lolos
4	Surya 12	1	25 %	Lolos
5	Aqua 1500 ml	1	25 %	Lolos
6	Sari Roti Krim Coklat	1	25 %	Lolos
7	Sampoerna 16 Mild	1	25 %	Lolos

3. Analisa Data Frequence 2-itemset

Analisa kombinasi frequence 2-itemset pada iterasi kedua dilanjutkan dengan perhitungan kandidat 2 *itemset* dan hitung nilai *support*-nya dilakukan analisa dilakukan dengan menggunakan kombinasi *I-itemset* yang lebih dari 25% dari data transaksi tersebut menghitung jumlah *support*-nya maka akan mendapatkan hasil kombinasi barang 2-itemset barang

**Tabel 3. Data Frequence I-Itemset**

No.	Atribut Barang 1	Atribut Barang 2	Jumlah	Support	Ket
1	Aqua 600 ml	Capucino	1	25 %	Lolos
2	Aqua 1500 ml	Sari Roti Krim Coklat	1	25 %	Lolos
3	Aqua 1500 ml	Sampoerna 16 Mild	1	25 %	Lolos
4	Sari Roti Krim Coklat	Sampoerna 16 Mild	1	25 %	Lolos

Setelah dilakukan Analisa sistem menentukan kombinasi *2-itemset* dari data transaksi tersebut menghitung jumlah *support*-nya Minimum *Support* yang ditentukan adalah 25% sehingga yang memiliki item-item kurang dari minimum *Support* yaitu 25% akan dihilangkan, untuk melanjutkan dalam pencarian itemset3 dengan mendapatkan hasil.

#### 4. Analisa Data Frequence 2-itemset

Analisa kombinasi frequence 3-itemset pada iterasi ketiga dilanjutkan dengan perhitungan kandidat 3-itemset dan hitung nilai *support*-nya 25% jika nilai support kurang dari 25% maka tidak akan dilakukan menggabungkan tiga atribut barang untuk mendapatkan hasil support dari sebuah itemset baru

**Tabel 4. Data Frequence 3-Itemset**

No.	Itemset	Support	Keterangan
1.	BRG0005 BRG0006 BRG0007	25 %	Lolos

Setelah Analisa sistem menentukan kombinasi *3-itemset* dari data transaksi tersebut menghitung jumlah *support*-nya Minimum *Support* yang ditentukan adalah 25% sehingga yang memiliki item-item kurang dari minimum *Support* yaitu 25% akan dihilangkan, untuk melanjutkan dalam pencarian aturan association rules untuk menentukan pola pembelian barang dengan mendapatkan hasil.

#### 5. Membentuk Aturan Association Rules

Analisa pembentukan aturan associatin rules pada tahap ini dilakukan setelah pola frekuensi tinggi ditemukan, barulah dicari aturan *Association Rules* yang memenuhi syarat untuk *confidence* dengan menghitung nilai *confidence* 50% jika nilai confidence kurang dari 50% maka tidak akan dilakukan aturan asosiatif sebagai berikut:

**Tabel 5. Aturan Association Rules**

No.	Kode Barang	Support	Confidence
1.	BRG0005, BRG0006	25 %	100 %
2.	BRG0005, BRG0007	25 %	100 %
3.	BRG0006, BRG0007	25 %	100 %
4.	BRG0007, BRG0006	25 %	100 %

Hasil perhitungan aturan association rules maka akan dibuatkan rules berdasarkan Tabel 5. untuk mendapatkan hasil association maka akan dideskripsi dalam bentuk rules berdasarkan data kombinasi 1 dan kombinasi 2, mendapatkan rules dengan nilai *confidence* 50% jika nilai *confidence* kurang dari 50% maka rules tidak dapat dibuatkan maka dapat dilihat sebagai berikut:

1. Jika membeli Aqua 1500 ml, Sari Roti Krim Coklat Kemungkinan 100 % akan membeli Sampoerna 16 Mild.

2. Jika membeli Aqua 1500 ml,Sampoerna 16 Mild Kemungkinan 100 % akan membeli Sari Roti Krim Coklat.
3. Jika membeli Sari Roti Krim Coklat,Sampoerna 16 Mild Kemungkinan 100 % akan membeli Aqua 1500 ml.
4. Jika membeli Sampoerna 16 Mild,Sari Roti Krim Coklat Kemungkinan 100 % akan membeli Aqua 1500 ml.

### **Kesimpulan**

Berdasarkan hasil penelitian ini membahas implementasi algoritma Apriori dalam sistem kecerdasan buatan untuk analisis pola pembelian di minimarket K92-Mart. Permasalahan utama yang dihadapi adalah seringnya barang yang dibutuhkan konsumen tidak tersedia karena kurangnya perhatian terhadap stok. Dengan menggunakan data mining dan algoritma Apriori, penelitian ini berhasil mengidentifikasi pola pembelian barang dengan menentukan hubungan asosiasi antar produk. Hasil analisis menunjukkan bahwa barang-barang tertentu sering dibeli bersamaan, sehingga dapat digunakan sebagai dasar untuk strategi penjualan yang lebih efektif. Implementasi algoritma Apriori dalam sistem ini membantu pengelola minimarket dalam mengambil keputusan yang lebih strategis, meningkatkan efisiensi manajemen stok, dan mengoptimalkan ketersediaan barang sesuai dengan pola permintaan pelanggan.

1. Penelitian ini melakukan pencarian frequent itemset pada tahap ini dilakukan analisa dilakukan dengan menggunakan kombinasi I-itemset dari data transaksi tersebut menghitung jumlah *support*-nya dengan min support 25% dan min confidence 50%.
2. Penelitian ini melakukan pencarian frekuensi tinggi ditemukan, barulah dicari aturan *Association Rules* yang memenuhi syarat untuk *confidence* dengan menghitung nilai *confidence* 50%.

### **Daftar Pustaka**

- Araf Aliwijaya, & Hanny Chairany Suyono. (2023). Peluang Pemanfaatan Big Data di Perpustakaan. *Info Bibliotheca: Jurnal Perpustakaan Dan Ilmu Informasi*, 4(2), 1–17. <https://doi.org/10.24036/ib.v4i2.397>
- Dzulkarnaen, R. (2020). Perancangan Aplikasi Data Mining Market Basket Analysis Pada Apotek Permata dengan Metode Hybrid-Dimension Association Rules. *Journal of Information Technology*, 2(2), 67–72. <https://doi.org/10.47292/joint.v2i2.35>
- Febriyanti, L., & Zakaria, H. (2023). Implementasi Data Mining Untuk Memprediksi Produktivitas Pada Tanaman Kacang Tanah Menggunakan Metode Naive Bayes (Studi Kasus : Perkebunan Kacang Tanah Di Kota Bogor). *LOGIC : Jurnal Ilmu Komputer Dan Pendidikan*, 1(2), 105–118.
- Katarina Sianturi, S., Satriani Fansuri, D., & Najmiatul Aini, W. (2023). Algoritma Apriori untuk Mengetahui Pola Beli Konsumen Pada Sistem Informasi Market Basket Analysis Berbasis Andriod. *Jurnal Insan Unggul*, 11(1), 35–58. <https://doi.org/10.47926/jiu.2023.11.1.35-58>
- Nofitri, R., & Eska, J. (2018). Implementasi Data Mining Klasifikasi C4.5 Dalam Menentukan Kelayakan Pengambilan Kredit. *Seminar Nasional Royal (SENAR)*.

- Royal, S. (2024). Perancangan Aplikasi Data Mining Untuk Menentukan Tingkat Kelarisan Produk Menggunakan Metode Clustering Dengan Algoritma K-Means. *Journal of Science and Social Research*, 4307(1), 116–123.
- Supardi, S., Karenina Ajie, A., Dwiyanti, A., Ramiaji, J., Jein, K., Aulia Ramadhanti, N., Maharani Putri, A., Ken Meilizar, R., & Penulis, K. (2023). Peran Data Mining dalam Memprediksi Tingkat Penjualan Sepatu Adidas Menggunakan Metode Algoritma Regresi Linear Sederhana. *Jurnal Ekonomi Manajemen Sistem Informasi*, 4(5), 883–890.
- Wismoaji, J. (2024). 1 ) *Teknik Informatika Politeknik Sawunggalih Aji*. 12(1), 36–49.