



ANALISIS FREKUENSI DATA MENGGUNAKAN ALGORITMA APRIORI (STUDI KASUS : TOKO AGEN SEMBAKO AAN)

¹Dicky Prasetya, ²Devi Damayanti

^{1,2}Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Pamulang, Kota Tangerang Selatan, Indonesia
Email: ¹dqprasetya916@gmail.com, ²dosen02390@unpam.ac.id

Abstrak

Jual beli adalah kegiatan umum manusia sebagai makhluk sosial yang memerlukan interaksi. Di Indonesia, banyak toko, baik barang maupun jasa, dapat diakses secara online, namun beberapa masih menggunakan perhitungan manual dengan kalkulator atau notepad. Data penjualan yang tidak terpakai dapat menjadi informasi strategis untuk masa depan. Dengan Algoritma Apriori, kita dapat menyimpulkan hasil dengan mencari kombinasi itemset yang sering dibeli, membentuk strategi sesuai kesimpulan penelitian. Keuntungan algoritma ini adalah kemampuannya bekerja dengan baik pada data besar atau sederhana, tergantung pada aplikasi dan metode yang digunakan. Dengan bantuan dari Bahasa pemrograman python yang sangat cocok untuk mengaplikasikan perhitungan dan dengan Desain dan operator dari RapidMiner yang tepat memastikan hasil sesuai dengan tujuan analisis data penjualan. Studi kasus di toko AAN menghasilkan beberapa temuan, seperti Telur Ayam dengan support tertinggi 26.67%, dan Chocolatos, Kapal Api Mix, Sardine ABC, dan Sunsilk dengan support minimum 10.00%. Kombinasi Indomie Goreng → Telur Ayam memiliki confidence tertinggi 100%, sementara kombinasi dua Indomie Kari Ayam → Telur Ayam memiliki confidence 87,5%.

Kata Kunci: Algoritma Apriori, Pertumbuhan Pola Asosiasi Tinggi, Python, RapidMiner, Data Penjualan

Abstract

Buying and selling is a common human activity as social beings that necessitates interaction. In Indonesia, numerous stores, offering both goods and services, can be accessed online, but some still rely on manual calculations using a calculator or notepad. Unused sales data can serve as strategic information for the future. With the Apriori Algorithm, we can derive conclusions by identifying frequently purchased itemset combinations, forming strategies based on research findings. The advantage of this algorithm lies in its ability to perform well with both large and simple datasets, depending on the application and methods employed. Assisted by the Python programming language, well-suited for implementing calculations, and with the design and operators from RapidMiner ensuring precise results aligned with the goals of sales data analysis. A case study in the AAN store yielded several findings, such as Chicken Eggs with the highest support at 26.67%, and Chocolatos, Kapal Api Mix, Sardine ABC, and Sunsilk with a minimum support of 10.00%. The combination of Indomie Goreng → Chicken Eggs has the highest confidence at 100%, while the combination of two Indomie Kari Ayam → Chicken Eggs has a confidence level of 87.5%.

Keywords: Aprior Algorithm, Frequent – Pattern Growth, Python, RapidMiner, Sales Data

1. PENDAHULUAN

Jual beli, sebagai aktivitas umum dalam interaksi sosial manusia, telah berkembang di Indonesia, Pertumbuhan jual beli sebagai bagian integral dari interaksi sosial manusia telah mengalami perkembangan pesat di Indonesia. Fenomena ini tidak hanya mencakup transaksi konvensional yang dilakukan secara langsung, tetapi juga merambah ke dunia daring melalui platform penjualan online. Namun, seiring dengan ekspansi ini, timbul beberapa tantangan, terutama terkait manajemen data penjualan yang seringkali tidak teroptimalkan. Fenomena ini dapat mengakibatkan ketidakseimbangan dalam persediaan barang, risiko kehilangan pelanggan, dan potensi pengurangan pendapatan bagi toko sembako AAN, disamping itu juga banyak data yang tidak terpakai dan akan menjadi tidak berguna jika tidak diolah sebagai sumber informasi dan strategi bisnis terbaik di kemudian hari.

Dalam rangka mengatasi permasalahan kompleks ini, penelitian ini mengarahkan fokusnya pada penerapan algoritma Apriori, sebuah algoritma data mining yang telah terbukti efektif, dalam konteks toko Agen sembako AAN. Algoritma Apriori, yang awalnya digunakan untuk analisis asosiasi dalam database transaksi, dapat menjadi solusi yang tepat untuk mengoptimalkan manajemen persediaan, mengidentifikasi pola pembelian, dan meramalkan permintaan pelanggan. Algoritma Apriori digunakan untuk menghasilkan aturan asosiasi dengan pola "if-then" atau "jika-maka", yang terfokus pada pembentukan keterkaitan antar elemen. Analisis keranjang belanja (Market Basket Analysis) merupakan salah satu teknik dalam ranah data mining yang mempelajari kebiasaan konsumen dalam melakukan pembelian produk secara bersamaan dalam suatu periode waktu tertentu [1].

Keputusan untuk mengeksplorasi algoritma Apriori tidak muncul begitu saja, melainkan sebagai respons terhadap tren positif penggunaannya oleh sejumlah pusat perbelanjaan terkemuka di luar negeri, seperti Walmart, serta penerapannya yang sukses di kalangan Usaha Mikro, Kecil, dan Menengah (UMKM) di Indonesia. Pada penelitian algoritma apriori menekankan tentang pentingnya data yang telah usang menjadi suatu informasi yang dihasilkan untuk mencari dan menerapkan strategi bisnis yang dapat mengoptimalkan, penjualan sehingga, tidak ada kerugian berlebih.

Dengan memanfaatkan data transaksi penjualan yang telah disimpan di dalam basis data, manajemen dapat memahami kebiasaan atau perilaku pelanggan terkait dengan buku-buku yang sering mereka beli. [2] Dengan melakukan analisis frekuensi data penjualan, penelitian ini berupaya memberikan kontribusi positif kepada pemilik toko, membantu mereka mengambil keputusan yang lebih tepat terkait penawaran diskon atau paket produk. Dengan menerapkan algoritma Apriori, penelitian ini berusaha memanfaatkan data penjualan untuk memahami preferensi konsumen dan mendukung pengambilan keputusan yang lebih baik terkait strategi pemasaran dan penargetan pasar [3].

Lebih dari sekadar pengoptimalan persediaan, penelitian ini berambisi untuk meminimalkan pemborosan dan mengoptimalkan penggunaan sumber daya informasi yang terdapat dalam data penjualan yang tidak terpakai. Dengan demikian, diharapkan bahwa hasil penelitian ini dapat membuka pintu untuk penerapan praktik terbaik dalam manajemen toko sembako, memajukan efisiensi operasional, dan meningkatkan daya saing dalam era bisnis yang terus berubah ini.

2. METODOLOGI PENELITIAN

2.1. Data Mining

“Data Mining adalah proses yang menggunakan metode matematika, statistika dan kecerdasan buatan untuk mengeksplorasi dan menemukan pola dan hubungan yang berarti dalam kumpulan data yang besar” [4]

Data mining adalah proses ekstraksi pola atau pengetahuan yang berguna dari suatu set data yang besar atau kompleks. Tujuan utama dari data mining adalah menemukan hubungan atau pola tersembunyi dalam data yang mungkin sulit atau tidak mungkin ditemukan secara manual [5].

2.2. Analisis Market Basket

metode statistik yang digunakan untuk mengidentifikasi pola pembelian konsumen dengan menganalisis item yang sering dibeli bersamaan dalam satu transaksi. Pendekatan ini umumnya digunakan dalam industri ritel dan e-commerce untuk memahami hubungan antarproduk dan membantu dalam strategi pemasaran, penataan produk, dan pengelolaan stok [6].

- a) Data Crawling
adalah proses ekstraksi informasi atau data dari berbagai sumber di web maupun pencarian secara real-time kepada data penelitian yang bersangkutan, berikut ini adalah contoh data yang diperoleh dari pemilik usaha dan data dibawah masih bersifat kotor atau belum di pre-processing.
- b) Data Cleaning (Preprocessing)
Data yang dihitung yaitu data grosir (data uji) dan data eceran (data testing) akan dibersihkan dengan cara memilah transaksi dengan menggunakan analisis market basket yang bisa menentukan transaksi mana yang sempurna dan ideal untuk pencarian dan perhitungan support dan confidence pada algoritma apiori.
- c) Data Transformation
Gunakan Bahasa pemrograman python dengan menggunakan beberapa library yang akan membantu dalam perhitungan, seperti numpy, pandas dll.
- d) Proses Mining
Proses mining menghitung secara manual menggunakan Microsoft Excel dengan rumus algoritma apiori mencari support tertinggi dan terendah beserta confidence tertinggi dan terendah.
- e) Evaluasi Pola
Mengidentifikasi pola support dan confidence yang akan ditarik kesimpulan dalam knowledge presentation, Dalam mengambil kesimpulan dalam penelitian ini mengandalkan visualisasi dari aplikasi RapidMiner.

2.3. Algoritma Apriori

apiori adalah algoritma ini dapat melakukan prediksi yang akurat dan dapat digunakan pada data dengan jumlah yang sangat besar serta dapat diimplementasikan secara mudah [7]. Sedangkan Metode Apriori merupakan pendekatan dalam pengumpulan data yang menggunakan aturan asosiasi (association rule) untuk mengidentifikasi keterkaitan asosiatif dari suatu kombinasi item [8].

2.4. Google Colab

Google colab adalah platform komputasi berbasis cloud yang dikembangkan oleh google. Platform ini dirancang untuk memudahkan menjalankan kode python dan menganalisis data di lingkungan terpisah dari perangkat local..Google Colab memungkinkan pengguna membuat, menggunakan dan berbagi notebook interaktif yang menggabungkan kode, teks dan visualisasi, lalu dipermudah dengan menggunakan eksekusi tulis, simpan dan bagikan [9].

2.5. Python

Python merupakan Bahasa pemrograman tingkat tinggi yang dirancang dengan mempertimbangkan keterbacaan kode. Bahasa ini memiliki sintaks yang sederhana dan lugas yang membuatnya mudah dipelajari dan digunakan untuk pengembangan baru dan berpengalaman. Python juga mendukung pemrograman fungsional, dan memiliki banya Pustaka dan kerangka kerja yang kuat. pemilihan Python sebagai bahasa pemrograman bertujuan untuk meningkatkan efisiensi dan kinerja eksekusi algoritma Apriori, sehingga proses analisis data dapat berjalan lebih cepat dan efektif [10].

3. HASIL DAN PEMBAHASAN

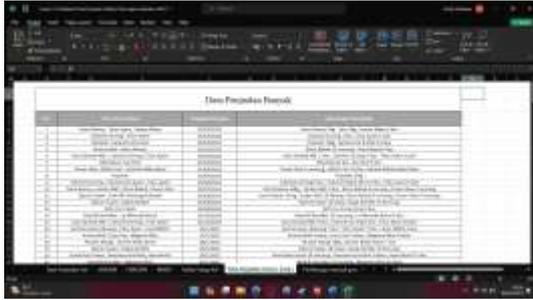
Pada Bagian ini menjelaskan tahapan-tahapan penelitian. Dimulai dengan perhitungan manual, Implementasi Python di Google Colab dan juga pembuktian hasil di aplikasi RapidMiner.

3.1. Perhitungan manual

Perhitungan manual menggunakan Microsoft Excel dengan cara menentukan minimum Support dan minimum Confidence dari data pada penelitian kali ini dengan menggunakan Teknik Analisis Market Basket. Total data yang didapat dari hasil crawling penelitian kali ini terdapat 214 data yang dibagi menjadi data grosir dan data satuan.

Gambar 1 Hasil Data Crawling

Data didapat dari wawancara kepada pemilik usaha yaitu ibu aan sendiri yang terdiri dari 3 bulan data yaitu Januari – Maret 2023 yang ditulis oleh ibu aan sendiri.



Gambar 2 Hasil Pembersihan Data

Setelah data dibagi menjadi 2 bagian maka mulailah perhitungan C1 – C3 pada perhitungan algoritma apriori yang bertujuan untuk mencari nilai minimum Support beserta Confidence nya.

Rumus mencari Support C1 :

$$Supp A = \frac{\text{Jumlah Transaksi Mengandung } A}{\text{Total Transaksi}}$$

Rumus mencari Support C2 :

$$Supp A \cap B = \frac{\text{Jumlah Transaksi Mengandung } A \text{ dan } B}{\text{Total Transaksi}}$$

Rumus mencari Confidence :

$$Confidence = P(B | A) = \frac{\text{Jumlah Transaksi Mengandung } A \text{ dan } B}{\text{Jumlah Transaksi Mengandung } A}$$

Dengan rumus dan perhitungan selanjutnya maka ditetapkan lah minimum supportnya **0,060** dan minimum confidencenya **0,3**. dari parameter yang telah ditetapkan maka didapat lah beberapa barang yang telah lulus minimum support sebagai berikut :

Table 1 Tabel Item Lolos Minimum Support

No	Nama Item	Support
1	Telur Ayam	0,267
2	Chocolatos	0,100
3	Kapal Api Mix	0,100
4	Sarden ABC	0,100
5	Sunsilk	0,100
6	Indomie Goreng	0,089
7	Indomie Kari Ayam	0,089
8	Setra Ramos	0,089
9	Djarum Super	0,078
10	Head & Shoulder	0,078
11	Sampoerna Kretek	0,078
12	Magnum Blue	0,067

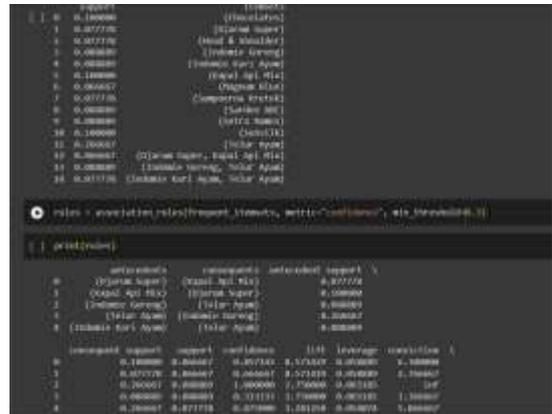
Data minimum support yang didapat adalah seperti ini, sedangkan data kombinasi minimum confidence adalah sebagai berikut :

Table 2 Tabel Kombinasi Lolos Minimum Confidence

No	Kombinasi	Supp	Conf
1	Kapal Api Mix → Djarum Super	0,089	1
2	Djarum Super → Kapal Api Mix	0,078	0,875
3	Indomie Kari Ayam → Telur Ayam	0,067	0,667
4	Indomie Goreng → Telur Ayam	0,067	0,857

3.2 Implementasi Python

Untuk perhitungan manual awalnya data yang berbentuk table dibuat menjadi array terlebih dahulu, agar memungkinkan bisa dimasukkan kedalam codingan python setelah itu hasil dari codingan python tersebut diatur parameter minimum support dan minimum confidence yang didapat dari perhitungan manual sebelumnya. Maka hasilnya adalah sebagai berikut :



Gambar 3 Hasil Implementasi Python pada Data Transaksi

Setelah data di implementasikan kedalam python, maka dengan memakai cara ini maka tugas python sebagai pengolah data otomatis yang bisa melengkapi perhitungan manual sebelumnya telah selesai.

3.3 Hasil Model RapidMiner

Selanjutnya pada tahap selanjutnya akan dilakukan pelatihan model visualisasi perhitungan terhadap aplikasi RapidMiner agar bisa terlihat seperti apa hubungan antar item dan kombinasi-kombinasinya.



Gambar 4 Model Penelitian RapidMiner

Model yang akan diuji terdiri dari 3 operator yaitu Numerical to Binominal, FP-Growth dan Association rules.

Untuk hasil yang dihasilkan dari model diatas setelah menerapkan Batasan-batasan parameter berupa minimum support dan minimum confidence yaitu seperti ini :



Gambar 5 Hasil Perhitungan RapidMiner

Dan untuk hasil kombinasi dari confidence adalah seperti gambar berikut :

Association Rules

```
Association Rules
{Telur Ayam} --> {Indomie Goreng} (confidence: 0.133)
{Kapal Api Mix} --> {Djarum Super} (confidence: 0.467)
{Djarum Super} --> {Kapal Api Mix} (confidence: 0.957)
{Indomie Kari Ayam} --> {Telur Ayam} (confidence: 0.875)
{Indomie Goreng} --> {Telur Ayam} (confidence: 1.000)
```

Gambar 6 Hasil Perhitungan Confidence RapidMiner

Hasil Minimum confidence yang diterapkan merupakan 0.3 maka hasil yang ditunjukkan adalah seperti gambar diatas. Untuk hasil visualisasi nya akan ditunjukkan graph seperti gambar dibawah ini



Gambar 7 Hasil Visualisasi RapidMiner

Visualisasi dibutuhkan untuk Knowledge Presentation dan untuk memaparkan hasil data dengan lebih mudah.

4. KESIMPULAN

Setelah melakukan penelitian lebih lanjut tentang algoritma apriori maka dapat disimpulkan algoritma apriori mengelompokkan item berdasarkan item yang sering dibeli berdasarkan minimum support dan minimum confidence, support dihitung dari seberapa sering kumpulan item muncul dalam suatu data, sedangkan untuk confidence didapat dari seberapa sering suatu aturan asosiasi terjadi dan dihitung benar. Contoh dalam kumpulan item support yaitu pada item **Telur Ayam** yang mempunyai minimum support terbesar yaitu sebesar **26.67%**, lalu ada **Chocolatos, Kapal Api Mix, Sarden ABC dan Sunsilk** memiliki minimum support sebesar **10.00%**. Untuk Nilai Confidence terbesar dimiliki oleh kombinasi item **Indomie Goreng** → **Telur Ayam** sebesar **1**. Selanjutnya ada kombinasi dua item dari **Indomie Kari Ayam** → **Telur Ayam** memiliki Confidence sebesar **0.875** atau **87.5%**.

5. REFERENCES

- [1] A. Rifqy Alfian, A. Hafidzul Kahfi, M. Rizky Kusumayudha, and M. Rezki, "Analisis Market Basket Dengan Algoritma Apriori Pada Transaksi Penjualan Di Freshfood," *IJCIT (Indonesian Journal on Computer and Information Technology*, vol. 4, no. 1, pp. 1–8, 2019.
- [2] A. Aquila, C. Pabendon, and H. D. Purnomo, "JURNAL MEDIA INFORMATIKA BUDIDARMA Penerapan Algoritma Apriori dan FP-Growth Untuk Market Basket Analisis Pada Data Transaksi NonPromo," 2023, doi: 10.30865/mib.v7i3.6153.
- [3] A. Hero Wilarto and P. L. Frans S, "PENENTUAN MINAT KONSUMEN TERHADAP PRODUK MENGGUNAKAN ALGORITMA APRIORI PADA PT. TELKOM INDONESIA," 2022. [Online]. Available: <https://ejournal.teknokrat.ac.id/index.php/teknoinfo/index>
- [4] Muhammad Alwi, Ninis Anggraini, and Rodia, "ANALISIS DATA MINING PADA PEMILIHAN JENIS GAME TERPOPULER MENGGUNAKAN ALGORITMA APRIORI," *Jurnal Teknoif Teknik Informatika Institut Teknologi Padang*, vol. 11, no. 1, pp. 9–15, Apr. 2023, doi: 10.21063/jtif.2023.v11.1.9-15.
- [5] S. Universitas Buana Perjuangan Karawang, A. Fauzi Universitas Buana Perjuangan Karawang, I. afauzi, and ubpkarawang Candra Zonyfar, "Data Mining Penerapan Algoritma Apriori untuk Analisis Data Transaksi Sistem Inventory (Studi Kasus PT ABC President Indonesia)," vol. III, no. 1, p. 143, 2022.
- [6] I. Qoni'ah and A. T. Priandika, "Analisis Market Basket Untuk Menentukan Asosiasi Rule Dengan Algoritma Apriori (Studi Kasus: Tb. Menara)," *Jurnal Teknologi dan Sistem Informasi*, vol. 1, no. 2, pp. 26–33, 2020, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI>
- [7] Z. Abidin, A. Kharisma Amartya, and A. Nurdin, "PENERAPAN ALGORITMA APRIORI PADA PENJUALAN SUKUCADANG KENDARAAN RODA DUA (STUDI KASUS: TOKO PRIMA MOTOR SIDOMULYO)," 2022. [Online]. Available: <https://ejournal.teknokrat.ac.id/index.php/teknoinfo/index>
- [8] W. Purba, "PENERAPAN DATA MINING MENGGUNAKAN METODE APRIORI UNTUK MENYUSUN POLA PERSEDIAAN INVENTARIS BARANG PADA BIRO SARANA DAN PRASARANA UNIVERSITAS PRIMA INDONESIA."
- [9] J. Homepage *et al.*, "IJIRSE: Indonesian Journal of Informatic Research and Software Engineering Naïve Bayes Classifier (NBC) Algorithm Analysis for Prediction Medical Device Sales," vol. 3, no. 2, pp. 119–126, 2023.
- [10] I. Wayan, P. Pratama, and O. Asroni, "Meningkatkan Wawasan Bisnis dan Strategi Pemasaran pada Toko Deli Point Melalui History Data Transaksi," vol. 4, no. 3, p. 163, doi: 10.36596/jpkmi.v4i3.669.