

Penerapan Data Mining Untuk Menganalisis Penjualan Produk Menggunakan Algoritma Apriori Berbasis WEB

Syarief Afifi Sumantri

Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Negeri Medan
E-mail: syariefafifi@gmail.com

Hermawan Syahputra

Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Negeri Medan
E-mail: hsyahputra@unimed.com

Jalan Willem Iskandar Pasar V Medan Estate, Kotak Pos No. 1589 Medan 20221 A,
Sumatera Utara

Abstract. *This study aims to determine the best selling food and beverage products at Caffe Kopi Kito. Data mining is the process of extracting useful information and patterns from very large data. Data mining includes data collection, data extraction, data analysis, and data statistics. The Apriori algorithm is a classic algorithm in data mining. This algorithm is used to see the intensity of occurrence of the relevant itemset or frequent items or association rules. This study uses consumer transaction data for 30 days in January 2023. Transaction data will be collected first based on the day and number of transactions, then the transaction data that has been collected will be grouped according to each item, the data that has been grouped will be carried out a priori algorithm process to determine the most dominant product. Then a system design will be carried out whose result will be a website. The results showed that using the website-based a priori algorithm could determine the most dominant product at Caffe Kopi Kito and make it easier for users to determine the most dominant product. Based on the results of product sales analysis at Caffe Kopi Kito, it can be concluded that working on the a priori algorithm on Caffe Kopi Kito using a website can be said to have the result of a product combination and in the future it can be used to create the best-selling menu packages at Caffe Kopi Kito.*

Keywords: *Data Mining, Apriori Algorithms, Transaction Data, Websites.*

Abstrak. Penelitian ini bertujuan untuk menentukan produk makanan dan minuman yang paling laris di Caffe Kopi Kito. Data mining merupakan proses penggalian informasi dan pola yang bermanfaat dari data yang sangat besar. Data mining mencakup pengumpulan data, ekstraksi data, analisis data, dan statistik data. Algoritma Apriori merupakan algoritma klasik dalam data mining. Algoritma ini digunakan untuk melihat intensitas kemunculan itemset atau frequent item atau aturan asosiasi yang relevan. Penelitian ini menggunakan data transaksi konsumen selama 30 hari pada bulan Januari 2023. Data transaksi akan dikumpulkan terlebih dahulu berdasarkan hari dan banyaknya transaksi, kemudian data transaksi yang telah di kumpulkan akan di kelompokkan sesuai dengan item masing-masing, data yang telah di kelompokkan akan dilakukan proses algoritma apriori

Received Mei 30, 2023; Revised Juni 30, 2023; Accepted Juli 15, 2023

* Syarief Afifi Sumantri, syariefafifi@gmail.com

untuk menentukan produk yang paling dominan. Kemudian akan dilakukan perancangan sistem yang hasilnya akan berupa website. Hasil penelitian menunjukkan dengan menggunakan algoritma apriori dengan berbasis website dapat menentukan produk yang paling dominan pada Caffe Kopi Kito dan mempermudah pengguna dalam menentukan produk yang paling dominan. Berdasarkan hasil analisis penjualan produk pada Caffe Kopi Kito dapat disimpulkan bahwa pengerjaan algoritma apriori pada Caffe Kopi Kito dengan menggunakan website dapat dikatakan memiliki hasil kombinasi produk dan untuk kedepannya bisa digunakan untuk membuat paket menu terlaris pada Caffe Kopi Kito.

Kata kunci: Data Mining, Algoritma Apriori, data transaksi, Website

LATAR BELAKANG

Pertumbuhan industri makanan dan minuman mulai menjadi andalan sebagai penopang pertumbuhan ekonomi nasional dan manufaktur. Pada tahun 2019 sektor makanan dan minuman terus menjadi primadona terutama pada sektor industri pariwisata yang mengembangkan wisata kuliner. Kementerian mencatat sektor makanan dan minuman memiliki nilai investasi pada tahun 2018 senilai Rp56,20 triliun. Pada proyeksi pertumbuhan nonmigas 2019, sektor industri makanan dan minuman diproyeksikan akan tumbuh tinggi sebesar 9,86%

Perkembangan teknologi informasi sangatlah besar memberikan kontribusi pada pertumbuhan jumlah data yang dikumpulkan dan disimpan dalam basis data. Maka dari itu dibutuhkan sebuah metode atau teknik yang dapat merubah kumpulan data tersebut menjadi sebuah informasi berharga ataupun pengetahuan yang bermanfaat untuk mendukung pengambilan keputusan bisnis. Dalam berbisnis tentunya ingin mencari keuntungan yang merupakan tujuan dasar dari terbentuknya unit bisnis. Maka untuk memenuhi tujuan tersebut diperlukan suatu strategi yang dapat meningkatkan penjualan.

Tujuan penggunaan algoritma apriori adalah untuk dapat dihasilkan informasi baru seperti pengetahuan tingkat kecenderungan atau kebiasaan pelanggan dalam membeli barang-barang secara bersamaan, menentukan tata letak produk, dan menentukan antrian produk-produk yang akan diproduksi lebih dulu, yang berguna untuk meningkatkan kinerja operasional serta sebagai referensi marketing Kopi Kito untuk meningkatkan keuntungan dari usaha kaki lima ini.

Dengan berkembangnya usaha Kopi Kito membuat jumlah data yang begitu besar justru bisa menjadi masalah bagi Kopi Kito tersebut jika tidak bisa dimanfaatkan sebaik mungkin. Semakin banyak data, maka Kopi Kito tersebut semakin memerlukan

usaha untuk mengolah data-data agar dapat dijadikan informasi yang berguna di masa depan, yang bisa digunakan untuk mengambil suatu keputusan dan untuk memperoleh pengetahuan yang baru (knowledge), yang berguna untuk peningkatan kemajuan suatu usaha, seperti peningkatan penjualan produk, penentuan produk yang akan diproduksi terlebih dahulu, promosi dll. Oleh karena itu sangat diperlukan sebuah aplikasi yang bisa memanfaatkan kumpulan data yang besar, agar dapat diperoleh informasi yang berguna bagi pengguna di masa depan.

Rumusan Masalah

Pengolahan data transaksi penjualan pada kafe Kopi Kito masih dilakukan secara manual, sehingga mengakibatkan kesulitan dalam menemukan item makanan dan minuman yang paling sering dibeli oleh konsumen guna membuat rekomendasi paket menu.

Batasan Masalah

Berdasarkan rumusan masalah dan tujuan diatas, maka pembatas dari penelitian ini adalah :

1. Data yang digunakan adalah data transaksi penjualan konsumen pada bulan Januari 2023.
2. Algoritma yang digunakan untuk membuat aturan asosiasi adalah algoritma apriori.

KAJIAN TEORITIS

Data Mining

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. Analisis asosiasi atau association rule mining adalah teknik data mining untuk menentukan aturan asosiasi antara satu kombinasi item. Interestingness measure yang dapat digunakan dalam data mining adalah (Khairul 2015).

- a. Support adalah suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu item atau itemset dari keseluruhan data transaksi.
- b. Confidence adalah suatu ukuran yang menunjukkan hubungan antara dua item secara conditional.

Tahap-Tahap Data Mining

Berikut ini adalah tahapan-tahapan data mining sebagai berikut:

a. Pembersihan Data (Data Cleaning)

Merupakan proses menghilangkan noise dan data yang tidak relevan. Pada umumnya data yang diperoleh, baik dari database suatu perusahaan maupun hasil eksperimen, memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak valid atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut-atribut data yang tidak relevan dengan hipotesa data mining yang dimiliki. Data-data yang tidak relevan itu juga lebih baik dibuang. Pembersihan data juga akan mempengaruhi performansi dari teknik data mining karena data yang ditangani akan berkurang jumlah dan kompleksitasnya

b. Integrasi Data (Data Integration)

Merupakan penggabungan data dari berbagai database baru. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dan lainnya. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya.

c. Seleksi Data (Data Selection)

Data yang ada pada database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang akan diambil dari database. Sebagai contoh, sebuah kasus yang meneliti faktor kecenderungan orang membeli dalam kasus market basket analysis, tidak perlu mengambil nama pelanggan, cukup dengan id pelanggan saja

d. Data Transformation

Data diubah atau digabungkan kedalam format yang sesuai untuk diproses dalam data mining. Beberapa metode data mining membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa metode standar seperti analisis asosiasi dan clustering hanya bisa menerima input data kategorikal. Karenanya data berupa angka numerik yang berlanjut perlu dibagi menjadi beberapa interval.

e. Proses Mining

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

f. Evaluasi Pola (Pattern Evaluation)

Untuk mengidentifikasi pola-pola menarik kedalam knowledge based yang ditemukan bermanfaat. Dalam tahap ini hasil dari teknik data mining berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti menjadikannya umpan balik untuk memperbaiki proses data mining, mencoba metode data mining lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat.

g. Presentasi Pengetahuan

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil data mining.

Algoritma Apriori

Algoritma Apriori merupakan suatu algoritma yang diajukan oleh R. Agrawal dan R. Srikant pada tahun 1994 pada suatu penelitian yang berguna untuk menemukan *frequent itemset* untuk *Boolean association rule*. Nama Algoritma Apriori berdasarkan fakta bahwa algoritma tersebut menggunakan pengetahuan sebelumnya (*prior knowledge*) dari pencarian *frequent itemset mining* (Gama 2016).

Aturan Asosiasi atau disebut dengan *association rules* merupakan salah satu teknik data mining yang berguna dalam mencari aturan asosiatif dari suatu barang. Aturan Asosiasi dibentuk dengan cara menganalisis pola data yang sering muncul (*frequent pattern*) dan dengan menggunakan parameter *support* dan *confidence* untuk mengidentifikasi hubungan yang paling penting dan sesuai. *Support* adalah indikasi yang menunjukkan seberapa sering suatu *item* muncul di dalam suatu *database*. Sedangkan *Confidence* menunjukkan berapa kali pernyataan tersebut benar (Gunadi 2016). Menurut Kusri (2017) Analisis asosiasi didefinisikan sebagai suatu proses untuk menemukan semua aturan asosiasi yang memenuhi syarat *minimum* untuk *support* (*minimum support*) dan syarat *minimum* untuk *confidence* (*minimum confidence*). *Minimal Support* : sebuah nilai yang ditentukan sendiri oleh peneliti untuk memangkas kombinasi *item set* menjadi lebih sedikit. Biasanya nilai ini dilihat dari kemunculan *item set* dalam data. *Minimal Confidence* sebuah nilai yang ditentukan juga oleh peneliti untuk memangkas kombinasi setiap *k-item set* (hasil dari pemangkasan *minimal support*) untuk membentuk aturan

asosiasi. Nilai ini ditentukan untuk mendapatkan kepastian yang tinggi terhadap kombinasi *item* (Desti Fitriati, 2016).

Berikut merupakan penjelasan mengenai *support* dan *confidence* pada aturan asosiasi:

a. *Support*

Support adalah suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu *item/item set* dari keseluruhan transaksi. Ukuran ini akan menentukan apakah suatu *item/item set* layak untuk dicari *confidence* selanjutnya atau tidak. Nilai *support* sebuah *item* diperoleh dengan menggunakan rumus sebagai berikut:

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

Sedangkan nilai *support* dari 2 *item* diperoleh dengan menggunakan rumus:

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

b. *Confidence*

Confidence adalah suatu ukuran yang menunjukkan hubungan antar 2 (dua) *item* secara *conditional* seperti: seberapa sering *item* A dibeli jika orang membeli *item* B. Nilai *confidence* dari aturan dari $A \cap B$ diperoleh dengan rumus berikut:

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

METODE PENELITIAN

Jenis penelitian yang digunakan dalam penelitian ini ialah penelitian kuantitatif yang menggunakan jenis data sekunder yang datanya diperoleh peneliti atau pengumpul data secara tidak langsung. Dikatakan tidak langsung karena data diperoleh melalui perantara, yaitu lewat orang lain, ataupun lewat dokumen. Adapun data yang diperoleh dari Kopi Kito ialah data transaksi makanan dan minuman selama 1 bulan pada bulan Januari 2023. Langkah-langkah yang digunakan dalam penelitian ini ialah:

1. Melakukan observasi pada Kopi Kito.
2. Mengumpulkan data transaksi makanan dan minuman dari Kopi Kito selama 1 bulan pada bulan Januari 2023
3. Pengolahan data mining:
 - a. Seleksi data.
 - b. *Preprocessing / cleaning data*.
 - c. *Transforming*.

- d. Data Mining.
- e. *Interpretation / Evaluation*.
- f. Menarik kesimpulan.

HASIL DAN PEMBAHASAN

Analisa

Analisa data dengan menggunakan algoritma apriori dapat dilakukan dengan menggunakan beberapa tahap, dimulai dengan pengelompokan data transaksi. Data yang diambil merupakan data transaksi konsumen pada bulan Januari tahun 2023. Data tersebut adalah data yang mewakili keseluruhan data transaksi sebanyak 8 transaksi dan dapat dilihat pada Tabel berikut.

Tabel 1. Data Transaksi

Transaksi	Item Yang Dibeli
1	Teh Manis, Nasi Goreng
2	Ayam Penyet, Sanger, Nasi Goreng, Teh Manis
3	Teh Manis, Ayam Penyet, Nasi Goreng, Thai Tea
4	Thai Tea, Ayam Penyet, Sanger
5	Nasi Goreng, Teh Manis, Sanger, Ayam Penyet
6	Teh Manis, Ayam Penyet, Nasi Goreng
7	Teh Manis, Ayam Penyet, Thai Tea
8	Nasi Goreng, Thai Tea

Aturan asosiasi yang berurut berdasarkan *minimum support* dan *minimum confidence* dapat dilihat pada tabel 2.

Tabel 2. Aturan Asosiasi Final

Itemset	Banyak Item	Support	Confidence
Nasi Goreng, Ayam Penyet	4	50%	67%
Nasi Goreng, Teh Manis	5	62,5%	83%
Ayam Penyet, Teh Manis	5	62,5%	83%
Teh Manis, Ayam Penyet, Nasi Goreng	4	50%	80%

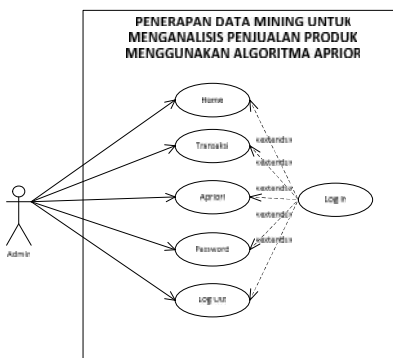
Dari hasil di atas didapat kesimpulan yang harus dilakukan oleh pihak kafe ialah menyediakan stok nasi goreng/bahan-bahan pembuatan dan stok ayam penyet dalam jumlah yang sama, menyediakan stok ayam penyet dan teh manis dalam jumlah yang sama, menyediakan stok nasi goreng dan teh manis dalam jumlah yang sama, dan juga pihak kafe juga bisa membuat keputusan untuk membuat paket makanan seperti paket promosi kombo untuk menambah daya tarik konsumen.

Perancangan

1. Use Case

Use case diagram adalah satu dari berbagai jenis diagram UML (*Unified Modelling Language*) yang menggambarkan hubungan interaksi antara sistem dan aktor. *Use Case* dapat mendeskripsikan tipe interaksi antara si pengguna sistem dengan sistemnya.

Use case diagram pada desain proses ini adalah untuk menggambarkan proses dari admin, selaku aktor yang berperan dalam penggunaan sistem sebagai pengelola data.



Gambar 1. Use Case Diagram

Implementasi Sistem

1. Implementasi

Implementasi sistem merupakan hasil dari pembuatan sistem yang sudah dirancang. Hasil implementasi berupa tampilan - tampilan dari sistem yang dibangun.

a. Menu Login

Halaman awal ini merupakan halaman login yang pertama kali admin lihat saat membuka website. Admin harus memasukkan username dan password yang sudah didaftarkan, jika yang dimasukkan salah maka admin tidak akan bisa masuk ke halaman selanjutnya. Jika berhasil terverifikasi maka admin akan dibawa menuju halaman dashboard. Jika tidak maka akan muncul pop up bahwa username dan password yang dimasukkan tidak sesuai atau login gagal.



Gambar 2. Menu Login

b. Menu Dashboard

Dashboard merupakan halaman pertama yang dapat ditemukan setelah berhasil melakukan login, terdapat menu-menu yang dapat dipilih untuk menjalankan sistem algoritma apriori pada data transaksi.



Gambar 3. Menu Dashboard

c. Menu Data Transaksi

Pada halaman menu data transaksi, admin dapat menampilkan data transaksi. Pada halaman ini, admin juga dapat menjalankan fungsi tambah (input), cari (search), edit (update) dan hapus (delete) data transaksi.

No	Item	Qty	Price	Total
1	teh manis	1	1000	1000
2	teh manis	1	1000	1000
3	teh manis	1	1000	1000
4	teh manis	1	1000	1000
5	teh manis	1	1000	1000
6	teh manis	1	1000	1000
7	teh manis	1	1000	1000
8	teh manis	1	1000	1000
9	teh manis	1	1000	1000
10	teh manis	1	1000	1000

Gambar 4. Menu Data Transaksi

d. Menu Proses Apriori

Pada halaman menu proses apriori, admin dapat melakukan proses apriori. Pada halaman ini, admin juga dapat menjalankan fungsi *min support*, *min confidence* dan hitung.

No	Item	Qty	Price	Total
1	teh manis	1	1000	1000
2	teh manis	1	1000	1000
3	teh manis	1	1000	1000
4	teh manis	1	1000	1000
5	teh manis	1	1000	1000
6	teh manis	1	1000	1000
7	teh manis	1	1000	1000
8	teh manis	1	1000	1000
9	teh manis	1	1000	1000
10	teh manis	1	1000	1000

Gambar 5. Menu Proses Apriori

e. Menu Tampilan Hasil

Pada halaman menu tampilan hasil, admin dapat melihat seluruh hasil dan kesimpulan dari proses apriori yang telah dilakukan.

No	Rule	Support	Confidence	Lift Ratio	Lift Ratio
1	teh manis + nasi goreng > teh manis + nasi goreng	11,64%	91,67%	15,67	1,65
2	teh manis + nasi goreng > teh manis + nasi goreng	11,64%	91,67%	15,67	1,65

Gambar 6. Menu Tampilan Hasil

Penjelasan

1. *Support* 11,64%, artinya 11,64% dari semua transaksi yang dianalisis menunjukkan bahwa Teh manis dan Nasi goreng dibeli bersamaan, sedangkan *confidence* sebesar 91,67% menyatakan tingkat kepercayaan atau dapat dikatakan jika seorang konsumen membeli Teh manis maka terdapat 91,67% kemungkinan dia akan membeli Nasi goreng juga.

2. *Support* 14,81%, artinya 14,81% dari semua transaksi yang dianalisis menunjukkan bahwa Teh manis dan Ayam penyet dibeli bersamaan, sedangkan *confidence* sebesar 84,85% menyatakan tingkat kepercayaan atau dapat dikatakan jika seorang konsumen membeli Teh manis maka terdapat 84,85% kemungkinan dia akan membeli Ayam penyet juga.

Dari penjelasan di atas dapat disimpulkan bagi produsen Kafe Kopi Kito agar menyediakan stok teh manis, nasi goreng dan ayam penyet berupa bahan-bahan pembuatan dalam jumlah yang sama. Dan juga pihak Kafe Kopi Kito bisa membuat paket makanan seperti paket promosi kombo untuk menambah daya tarik konsumen.

KESIMPULAN DAN SARAN

Berdasarkan hasil pembahasan dan analisis maka disimpulkan bahwa data transaksi Kafe Kopi Kito yang dilakukan selama 1 bulan pada bulan Januari 2023 sebanyak 464 transaksi memiliki kombinasi tinggi terhadap teh manis, nasi goreng dan teh manis, ayam penyet. Dan juga pihak Kafe Kopi Kito bisa membuat paket makanan seperti paket promosi kombo untuk menambah daya tarik konsumen. Bagi peneliti kedepannya agar melakukan pengembangan terhadap aplikasi ini agar input transaksi dapat dilakukan oleh konsumen sendiri serta dilakukan pengembangan terhadap website ini agar dapat menampilkan grafik analisa pemesanan.

DAFTAR REFERENSI

- Ade, H., (2016): Pemodelan UML Sistem Informasi Monitoring Penjualan dan Stok Barang (Studi Kasus: Distro Zhezha Pontianak, *Jurnal Khatulistiwa Informatika*, 4(2).
- Agus, N., T. H., (2016): Implementasi Algoritma Apriori Untuk Analisa Penjualan Dengan Berbasis Web, *Jurnal Simetris*, 7(2). 701-706.
- Desak, M, D, U, P., S. B., (2017): Penerapan Data Mining Pada Penjualan TAN'S BAKERY Menggunakan Algoritma Apriori, *Jurnal TIK*, 3(2). 164-174.
- Desti, F. (2016): Implementasi Data Mining untuk Menentukan Kombinasi Media Promosi barang Berdasarkan Perilaku Pembelian pelanggan Menggunakan Algoritma Apriori, *Annual Research Seminar* , 2(1). 473.
- Gama, A. W. O., Putra, I. K. G. D., & Bayupati, I. P. A. (2016). *Majalah ilmiah teknologi elektro. Majalah Ilmiah Teknologi Elektro* (Vol. 15).
- Gunadi, G., & Sensuse, D. I. (2016). Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori dan Frequent Pattern Growth (FP-Growth) : Studi Kasus Percetakan PT. GRAMEDIA. *Telematika MKOM*, 4(1), 118–132.
- Khairul U., (2015): Analisa Data Mining Dalam Penjualan Sparepart Mobil Dengan Menggunakan Metode Algoritma Apriori, *Jurnal Sistem Informasi*, ISSN : 2085–1367.
- Muhammad, F., F. A, S., (2019): Analisis Algoritma Apriori Pada Pemesanan Konsumen Di Café The L.CO Coffe, *Jurnal SAINTTEK*, 1(1). 52-57.
- Santosa, B. (2007). *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- Tika, D, A., T, I, H., I, K., (2017): Analisis Data Mining Menggunakan Algoritma Apriori Untuk Meningkatkan *Cross Selling* Dan *Up Selling* (Studi Kasus Rumah Makan Mas Nur Perwakarta), *Jurnal TIK*.