
PENERAPAN METODE ASOSIASI DATA MINING MENGGUNAKAN ALGORITMA APRIORI UNTUK MENGETAHUI KOMBINASI ANTAR ITEMSET PADA PONDOK KOPI

Fitri Nurchalifatun

Fakultas Ilmu Komputer, Universitas Dian Nuswantoro,
Jl. Nakula I no.5-11 Semarang, Indonesia
Telp. (024) 3517261. Fax : (024) 3520165
www.dinus.ac.id // Email: Fitrichalifa@gmail.com

ABSTRAK

Inventory merupakan salah satu aspek penting dalam suatu perusahaan. Lokasi Pondok Kopi yang sulit dijangkau membuat aktivitas distribusi bahan inventory menjadi kurang lancar. Untuk itu, sebelumnya perlu melakukan perencanaan inventory dengan baik, sebab inventory yang tidak mencukupi juga mempengaruhi tingkat penjualan. Untuk mengetahui kebutuhan inventory, dapat dilakukan analisis pasar untuk mencari tahu serangkaian item-item apa saja yang dibeli secara bersamaan oleh para konsumen melalui data transaksi. Dengan penerapan metode Market Basket Analysis (MBA) dan algoritma Apriori, dapat dilakukan analisis pada data transaksi sehingga menemukan informasi mengenai pola beli konsumen dan direpresentasikan dalam bentuk aturan asosiasi. Berdasarkan data transaksi penjualan pada Pondok Kopi, dilakukan analisis menggunakan algoritma apriori dengan parameter minimum support sebesar 5% dan minimum confidence sebesar 30%. Hasil dari penelitian tersebut menghasilkan 7 aturan asosiasi dengan kombinasi item terbesar sampai 2-itemset.

Kata kunci : Algoritma Apriori, Data Mining, Market Basket Analysis, Inventory

1. PENDAHULUAN

1.1 Latar Belakang

Pada saat ini dalam dunia bisnis semakin berkembang pesat, Khususnya dalam dunia perdagangan. Adanya persaingan yang semakin banyak dalam dunia bisnis, maka setiap pebisnis dituntut agar senantiasa dapat memikirkan cara-cara secara efektif dan efisien dalam pengembangan bisnis tersebut. Perusahaan bisnis dalam dunia hidangan merupakan perusahaan yang kegiatannya membeli bahan mentah dan menjual bahan yang siap saji untuk disajikan kepada konsumen, sehingga penjualan ini bertujuan untuk memperoleh keuntungan. Salah satu yang berhubungan dengan

pengembangan bisnis tersebut dengan diperlukannya pengelolaan dan pemeriksaan yang memadai terhadap inventori berbagai jenis stok bahan yang berada di gudang secara kontinu. pengelolaan inventori stok bahan merupakan fungsi manajerial yang sangat penting dalam mengelola inventori bahan yang dapat mempengaruhi efektifitas dan efisiensi kegiatan perusahaan.

Dengan melihat kebutuhan konsumen manajemen harus bisa memutuskan kapan akan melakukan pre-order pembelian pada suplier. Untuk itu dibutuhkan suatu sistem yang dapat membantu pihak pondok kopi untuk memberikan kendali yang handal untuk manajemen pondok kopi.

Pentingnya sistem persediaan bahan di suatu perusahaan dan jenis bahan yang harus di stok untuk mengantisipasi kekosongan bahan. Karena minimnya stok bahan yang di order serta lokasi yang susah di jangkau akan dapat berpengaruh pada inventory bahan, pelayanan konsumen dan pendapatan perusahaan.

Sistem yang dibutuhkan adalah sistem yang mendukung pengambilan keputusan, yang mana sistem dapat meningkatkan penjualan pada perusahaan dan memberikan saran untuk persediaan stok bahan mentah dari data transaksi. dengan sistem pendukung keputusan menggunakan aturan asosiasi berdasarkan algoritma apriori.

1.2 Rumusan Masalah

Berdasarkan latar belakang permasalahan yang di uraikan diatas, maka yang menjadi rumusan masalah dari penelitian ini adalah sebagai berikut :

1. Bagaimana menerapkan algoritma apriori dalam menentukan kombinasi antar itemset untuk membantu memprediksi inventory mendatang.
2. Bagaimana menerapkan metode MBA untuk menghasilkan *rule* dari pola kombinasi antar itemset.

1.3 Batasan Masalah

Agar tidak terlepas dari latar belakang dan perumusan masalah, maka penulis hanya membahas batasan masalah sebagai berikut :

1. Metode *data mining* yang digunakan adalah metode asosiasi (*market basket analysis*). Data yang digunakan adalah data transaksi penjualan item.
2. Menggunakan informasi data penjualan yang di peroleh berjumlah 108 data transaksi.

3. Objek penelitian ini adalah kafe/restoran.
4. Metode data mining yang digunakan adalah metode asosiasi (*Market Basket analysis*).
5. Algoritma yang digunakan dalam membentuk aturan asosiasi adalah algoritma apriori.

1.4 Tujuan dan Manfaat Penelitian

Berdasarkan perumusan masalah dan pembatasan masalah yang ada, maka tujuan yang ingin dicapai pada tugas akhir ini adalah menentukan aturan asosiasi menggunakan metode Market Basket Analysis dan algoritma Apriori berdasarkan data transaksi untuk mendapatkan informasi sebagai langkah pengambilan keputusan dalam memprediksi inventori untuk waktu yang akan datang. Serta manfaatnya yaitu memberikan pertimbangan untuk menentukan keputusan akhir dalam mengendalikan inventori bahan dan memberi pertimbangan dalam menyediakan inventori bahan untuk penjualan selanjutnya.

2. LANDASAN TEORI

2.1 Definisi Data Mining

Data mining atau yang sering disebut sebagai *knowledge discovery in database* (KDD) adalah kegiatan yang meliputi pengumpulan data, pemakaian data historis untuk menentukan keteraturan, pola atau hubungan dalam data berukuran besar. Keluaran *data mining* ini bisa dipakai untuk membantu pengambilan keputusan di masa mendatang (Subekti Mujiasih, "Pemanfaatan Data Mining Untuk Prakiraan Cuaca," 2011.)

2.2 Pengelompokan Data Mining

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan yaitu (Emha

Taufiq Luthfi and Kusri, *Algoritma Data Mining*. Yogyakarta: Andi Offset, 2009.):

1. Deskripsi
Terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Deskripsi dari pola kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola dan kecenderungan.
2. Klasifikasi
Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, pengelompokan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, pendapatan rendah.
3. Estimasi
Hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada kategori. Model ini dibangun menggunakan *record* lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.
4. Prediksi
Hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.
5. Pengklusteran
Kluster merupakan sekumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidak miripan dengan *record-record* dalam kluster lain. Pengklusteran berbeda dengan klasifikasi yaitu tidak adanya

variabel target dalam pengklusteran. Pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target. Akan tetapi, algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan *record* dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan *record* dalam kelompok lain akan bernilai minimal.

6. Asosiasi
Menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih sering disebut analisis keranjang belanja.

2.3 Metode Asosiasi

Metode asosiasi adalah metode yang menemukan suatu kombinasi item yang muncul bersamaan. Dalam dunia bisnis, sering disebut dengan *affinity Analysis* atau *Market Basket Analysis*. Aturan *Association rules* biasanya menggunakan pola “jika (*if*)” mewakili *antecedent* dan “maka (*then*)” mewakili *consequent*, bersamaan dengan pengukuran *support (coverage)* dan *confidence (accuracy)* yang terasosiasi dalam aturan. Fungsi *Association rules* sering kali disebut dengan “*Market Basket Analysis*”, yang digunakan untuk menentukan kombinasi diantara himpunan item-item. *Market Basket Analysis* adalah kebiasaan membeli konsumen dengan mencari asosiasi dan kombinasi antara item-item berbeda (Devi Dinda Setiawati).

2.4 Tahapan Association Rule

Metodologi dasar analisis asosiasi sebagai berikut :

- a. Analisis Pola Frekuensi Tinggi

Tahap ini mencari kombinasi antar item yang memenuhi syarat minimum dari nilai *support* dalam *database*.

Nilai *support* diperoleh dengan rumus :

$$Support(A) = \frac{\text{jumlah transaksi mengandung A}}{\text{total transaksi}} \times 100\%$$

Sementara nilai *support* dari 2-item diperoleh rumus :

$$Support(A,B) = \frac{\text{jumlah transaksi mengandung A dan B}}{\text{total transaksi}} \times 100\%$$

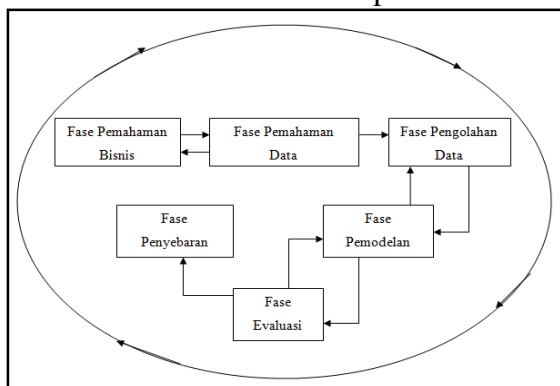
b. Pembentukan Aturan Asosiasi

Setelah semua pola frekuensi tertinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum *confidence* dengan menghitung *confidence* aturan asosiatif $A \rightarrow B$. Nilai aturan $A \rightarrow B$ diperoleh rumus :

$$Confidence(A \Rightarrow B) = P(A|B) = \frac{\text{jumlah transaksi mengandung A dan B}}{\text{jumlah transaksi mengandung B}} \times 100\%$$

2.5 CRISP-DM

CRISP-DM merupakan proses strategi dalam pemecahan masalah secara umum dari bisnis atau unit penelitian.



Gambar Proses CRISP-DM

Enam fase dalam CRISP-DM (Larose, 2005) :

1. Fase Pemahaman Bisnis (*Business Understanding Phase*)
 - i. Menentukan tujuan dan kebutuhan secara detail dalam

lingkup bisnis secara keseluruhan.

- ii. Menyiapkan strategi awal untuk mencapai tujuan.
- iii. Menerjemahkan tujuan dan batasan menjadi formula dari permasalahan *data mining*.

2. Fase Pemahaman Data (*Data Understanding Phase*)

- i. Pengumpulan data.
- ii. Mencari pengetahuan awal dan mengenali data lebih jauh dengan cara melakukan analisis penyelidikan data.
- iii. Memilih sebagian group data yang mungkin mengandung pola dari permasalahan.

3. Fase Pengolahan Data (*Data Preparation Phase*)

- i. Pilih kasus dan variabel yang ingin dianalisis.
- ii. Dilakukan perubahan pada variabel jika dibutuhkan.
- iii. Data awal disiapkan untuk perangkat pemodelan.

4. Fase Pemodelan (*Modeling Phase*)

- i. Pilih dan aplikasikan teknik pemodelan yang sesuai.
- ii. Kalibrasi aturan model untuk mengoptimalkan hasil.
- iii. Jika diperlukan, proses dapat kembali ke fase pengolahan data untuk menjadikan data ke dalam bentuk yang sesuai dengan spesifikasi kebutuhan teknik *data mining*.
- iv. Perlu di perhatikan bahwa beberapa teknik mungkin untuk digunakan pada permasalahan data mining yang sama.

5. Fase Evaluasi (*Evaluation Phase*)

- i. Menetapkan apakah terdapat model yang memenuhi tujuan pada fase awal.
- ii. Mengevaluasi satu atau lebih model yang digunakan dalam

fase permodelan untuk mendapatkan kualitas dan efektifitas sebelum disebarkan untuk digunakan.

- iii. Menentukan apakah terdapat permasalahan penting dari bisnis atau penelitian yang tidak di tangani dengan baik.
 - iv. Mengambil keputusan berkaitan dengan penggunaan hasil dari *data mining*.
6. Fase Penyebaran (*Deployment Phase*)
- i. Menggunakan model yang dihasilkan. Terbentuknya model tidak menandakan telah terselesainya proyek.
 - ii. Contoh sederhana penyebaran : pembuatan laporan.
 - iii. Contoh kompleks penyebaran : penerapan proses *data mining* secara paralel pada departemen lain.

2.6 Algoritma Apriori

Algoritma apriori merupakan jenis aturan asosiasi pada data mining yang digunakan untuk menentukan pola frekuensi tinggi. Pada tahun 1994 Agrawal dan Srikant mengusulkan suatu algoritma dasar untuk menentukan *frequent itemset* untuk aturan asosiasi yaitu Algoritma Apriori. Suatu asosiasi dikatakan penting atau tidak dapat diketahui dengan cara mencari nilai penunjang (*support*) dan nilai kepastian (*confidence*). Setelah menemukan *frequent itemset*, untuk menggali informasi maka algoritma kemudian meneliti *knowledge* dari *frequent item* sebelumnya (Yuliana Dewi, 2015). Analisis asosiasi biasanya didefinisikan sebagai proses untuk menemukan semua aturan asosiasi yang memenuhi syarat minimum untuk *support* dan syarat minimum untuk *confidence*.

3. METODOLOGI PENELITIAN

3.1 Objek Penelitian

Penelitian tugas akhir penulis melakukan penelitian pada data transaksi penjualan di Pondok Kopi yang berlokasi di Desa Sidomukti Kecamatan Bandungan Kabupaten Semarang.

3.2 Uraian Langkah Analisis Asosiasi

Mencari kombinasi pola item yang memenuhi syarat nilai minimum *support* dan *confidence*. Untuk mencari nilai *support* kandidat 1-itemset digunakan rumus :

$$Support(A) = \frac{\text{jumlah transaksi mengandung A}}{\text{total transaksi}} \times 100\%$$

Dimana nilai *support(A)* diperoleh dari membagi jumlah transaksi yang mengandung nilai A (satu item) dengan total transaksi di kali 100%.

Sedangkan untuk mencari nilai *support* untuk kandidat 2-itemset yaitu :

$$Support(A,B) = \frac{\text{jumlah transaksi mengandung A dan B}}{\text{total transaksi}} \times 100\%$$

Nilai *support (A,B)* diperoleh dengan membagi jumlah transaksi mengandung A dan B (dua item) dengan total seluruh transaksi dikalikan 100%.

Sedangkan untuk mencari aturan asosiasi dengan mencari nilai minimum *confidencenya* dengan rumus :

$$Confidence(A \Rightarrow B) = P(A|B) = \frac{\text{jumlah transaksi mengandung A dan B}}{\text{jumlah transaksi mengandung A}} \times 100\%$$

Nilai *confidence (A→B)* diperoleh dengan membagi jumlah transaksi mengandung A dan B (dua item) dengan jumlah transaksi

mengandung A dikalikan 100% sehingga aturan asosiasi akan terbentuk dengan nilai yang memenuhi syarat minimum *confidence* yang ditentukan.

4. Hasil Penelitian

Tabel hasil aturan asosiasi

No	Aturan asosiasi	support	confidence
1	roti umbul→vanila late	7,41%	47,06%
2	roti . umbul→tiramisu	5,56%	35,29%
3	tiramisu→vanila late	6,48%	33,33%
4	tiramisu→chokimisu	6,48%	33,33%
5	pisang keju→vanila late	5,56%	35,29%
6	chokimisu→tiramisu	6,48%	53,85%
7	kentang goreng→tiramisu	5,56%	37,50%

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

1. Dengan penerapan algoritma apriori dalam menentukan kombinasi antar itemset. Ditemukan 7 aturan asosiasi yaitu roti umbul → vanila late dengan *confidence* 47,06%, roti umbul → tiramisu dengan *confidence* 35,29%, tiramisu → vanila late dengan *confidence* 33,33%, tiramisu → chokimisu dengan *confidence* 33,33%, pisang keju → vanila late dengan *confidence* 35,29%, chokimisu → tiramisu dengan *confidence* 53,85%, dan yang terakhir kentang goreng → tiramisu dengan *confidence* 37,50%. Dari aturan yang dihasilkan dapat dijadikan sebagai informasi yang sangat berharga dalam pengambilan

keputusan untuk mengendalikan inventori mendatang.

2. Hasil pola kombinasi yang diperoleh dengan metode asosiasi yaitu yang mempunyai pola kombinasi *confidence* tertinggi adalah chokimisu maka tiramisu dengan nilai *confidence* 53,85%, roti umbul maka vanila late dengan nilai *confidence* 47,06%.

5.2 Saran

1. Disarankan agar dikembangkan dengan menambah jumlah data serta penggunaan *confidence* dan *support* yang bervariasi agar diperoleh hasil aturan asosiasi yang efisien.
2. Disarankan agar program aplikasi lebih dikembangkan lagi dengan aplikasi yang lain agar mudah dimengerti oleh pengguna.

6. Daftar Pustaka

- Dewi Yuliana, 2015, Penerapan Metode Market Basket Analysis dengan Algoritma Apriori Untuk Menentukan Kebijakan Promosi Pada Kedai Amarta.
- Kusrini dan Emha Taufiq Lutfhi, Algoritma Data Mining. Andi, Yogyakarta, 2009.
- Larose, Daniel T, 2005, Discovering Knowledge in Data: An Introduction to Data Mining. New Jersey : John Wiley & Sons.
- Mujiasih Subekti, 2011, Pemanfaatan Data Mining Untuk Prakiraan Cuaca.
- Setiawati Dewi Dinda, Penggunaan Metode Apriori Untuk Analisa Keranjang Pasar Pada Data Transaksi Penjualan Minimarket Menggunakan Java&MySQL.