

ANALISIS MARKET BASKET DENGAN ALGORITMA APRIORI (STUDY KASUS TOKO ALIEF)

Hernawati

STMIK Nusa Mandiri
Jl. Kramat Raya No. 18 Rt. 01/Rw. 07 Kwitang, Senen, Jakarta Pusat
watiherna27@gmail.com

ABSTRAK

Dalam persaingan dunia bisnis sekarang ini menurut para pelakunya untuk senantiasa mengembangkan bisnis mereka dan juga agar selalu bertahan dalam persaingan. Untuk mencapai hal itu, ada beberapa hal yang bisa dilakukan yaitu dengan meningkatkan kualitas produk dan penambahan jenis produk. Untuk memenuhi kebutuhan tersebut terdapat beberapa hal yang bisa dijalankan salah satunya dengan melakukan analisis data transaksi. Algoritma apriori termasuk jenis aturan asosiasi pada data mining. Aturan yang menyatakan asosiasi atau *association rule mining* adalah teknik data mining untuk menemukan aturan suatu kombinasi *item*. Salah satu tahap analisis asosiasi pola frekuensi tinggi (*frequent pattern mining*). Penting tidaknya suatu asosiasi dapat diketahui dengan dua tolak ukur, yaitu *support* dan *confidence*. Nilai penunjang (*support*) adalah persentase kombinasi *item* tersebut dalam database sedangkan nilai kepastian (*confidence*) adalah kuatnya hubungan antar-*item* dalam aturan asosiasi. Dari data 30 transaksi terdapat 1 pola *asosiasi* yang memenuhi syarat, salah satunya adalah jika membeli telur maka akan membeli rokok dengan nilai *confidence* tertinggi = 62,5%, sehingga membantu untuk mengambil keputusan perusahaan sebagai gambaran dalam rangka mendapatkan pola penjualan produk untuk promosi.

Kata Kunci : *Data mining, Algoritma Apriori, Support, Confidence.*

ABSTRACT

In today's business world competition according to the perpetrators to constantly develop their business and also to always survive in the competition. To achieve that, there are several things that can be done is to improve product quality and the addition of product types. To meet these needs there are several things that can be executed one of them by doing transaction data analysis. A priori algorithm includes the type of association rules in data mining. The rule that states association rule mining is a data mining technique to find the rules of a combination of items. One phase of analysis of the association of frequent pattern mining. The importance of an association can be identified by two benchmarks, namely support and confidence. Support value is the percentage of combination of items in the database while the value of certainty (confidence) is the strength of inter-item relationships in association rules. From 30 transaction data there is 1 pattern of eligible association, one of them is if buying eggs then will buy cigarettes with the highest confidence value = 62,5%, so help to take company decision as picture in order to get pattern of product sale for promotion.

Keywords: *Data mining, Apriori Algorithm, Support, Confidence.*

1. PENDAHULUAN

Dalam persaingan dunia bisnis sekarang ini menurut para pelakunya untuk senantiasa mengembangkan bisnis mereka dan juga agar selalu bertahan dalam persaingan. Untuk menacapai hal itu, ada beberapa hal yang bisa dilakukan yaitu dengan meningkatkan kualitas produk dan penambahan jenis produk. Untuk memenuhi kebutuhan tersebut terdapat beberapa hal yang bisa dijalankan salah satunya dengan melakukan analisis data transaksi.

Toko Alief merupakan toko kelontong yang bergerak dibidang penjualan kebutuhan pokok yang memiliki sistem seperti pada swalayan yaitu pembeli mengambil sendiri barang yang akan dibeli. Dari data penjualan pada toko alief selama ini tidak tersusun dengan baik, sehingga data tersebut hanya berfungsi sebagai arsip bagi toko dan tidak dapat dimanfaatkan untuk pengembangan stretegi pemasaran.

Dalam rangka menghadapi persaingan dalam pemasaran guna menghasilkan peningkatan pendapatan toko, pihak terkait mengambil keputusan

untuk menentukan strategi pemasaran produk yang akan dijual. Dengan data-data dijadikan sebagai sistem pengambilan keputusan untuk solusi bisnis serta dukungan infrastruktur dibidang teknologi yang merupakan penyebab munculnya suatu teknologi data mining.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan dataset transaksi atau market basket dengan jumlah 30 transaksi dengan berbagai macam item yang dibeli oleh konsumen. Dataset tersebut akan di olah untuk mencari *frequent itemset* menggunakan metode algoritma apriori kemudian dicari pola-pola belanja konsumen dengan *association rule* untuk mencari *support* dan *confidence* pada item yang saling berhubungan.

3. LANDASAN TEORI

Himpunan data penjualan yang dimiliki sebenarnya dapat diolah menggunakan data mining untuk melihat pola pembelian pelanggan, dengan data mining untuk data yang besar tidak akan terbuang begitu saja dan dapat bermanfaat memberikan keuntungan kepada perusahaan. Data mining dasarnya adalah untuk membangun program komputer yang menyaring *database* otomatis, mencari keteraturan atau pola. Pola yang kuat dan jika ditentukan adanya kemungkinan-kemungkinan akan menggeneralisasi untuk membuat prediksi yang akurat pada data masa depan. Akan tetapi akan ada masalah ketika banyak pola yang tidak menarik (Witten, dkk: 2011). Data mining adalah proses yang menggunakan teknik statistic, matematika, kecerdasan buatan dan *mechine learning* untuk mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database. Algoritma ini ditujukan untuk mencari kombinasi *item-set* yang mempunyai suatu nilai keseringan tertentu sesuai kriteria atau filter yang diinginkan (Kusrini: 2009).

Toko UcoK merupakan toko kelontong yang menjual sembako dalam pencatatan penjualannya hanya berupa arsip-arsip penjualan dan data transaksi dapat diolah dengan data mining sehingga data tersebut tidak hanya menjadi arsip, tetapi dapat memberikan pengetahuan tentang pola pikir konsumen, seperti tipe-tipe pembeli, hubungan antar item yang dibeli. Seperti tipe pembeli ibu rumah tangga ketika belanja kebutuhan, sebagian besar mereka membeli bahan pokok. Tetapi berbeda untuk tipe pembeli pelajar, ketika masuk tahun ajaran baru mereka akan membeli kebutuhan untuk belajar, dan sebagainya. Kesimpulan adalah market basket merupakan pengolahan data keranjang belanja konsumen. Contoh Ibu rumah tangga sering membeli ayam kemudian akan membeli telur, jarang sekali

setelah membeli ayam akan membeli kemeja dan pulpen (Dinata: 2017).

Pola yang sudah terbentuk pada keranjang belanja atau market basket dapat memberikan keputusan untuk display produk, promosi diskon produk, strategi penjualan untuk barang yang kurang menarik minat beli konsumen dan sebagainya, pola ini disebut dengan analisis asosiasi (*analysis association*). Kekuatan aturan asosiasi dapat diukur dengan *support* dan *confidence*. aturan asosiasi yang diambil adalah aturan-aturan asosiasi yang memiliki nilai *confidence* lebih besar atau sama dengan nilai minimum *confidence* yang telah ditentukan. Menggunakan hasil dari *support* dikalikan dengan *confidence* terhadap masing-masing aturan asosiasi yang dihasilkan untuk mengukur kekuatan dari aturan asosiasi yang bersangkutan (Hidayat: 2016).

Algoritma apriori adalah suatu algoritma dasar yang diusulkan oleh Agrawal & srikant pada tahun 1994 untuk penentuan *frequent itemset* untuk aturan asosiasi *boolean*. Algoritma apriori termasuk jenis aturan asosiasi pada data mining. Aturan yang menyatakan asosiasi atau *association rule mining* adalah teknik data mining untuk menemukan aturan suatu kombinasi *item*. Salah satu tahap analisis asosiasi pola frekuensi tinggi (*frequent pattern mining*). Penting tidaknya suatu asosiasi dapat diketahui dengan dua tolak ukur, yaitu *support* dan *confidence*. Nilai penunjang (*support*) adalah persentase kombinasi *item* tersebut dalam database sedangkan nilai kepastian (*confidence*) adalah kuatnya hubungan antar-*item* dalam aturan asosiasi (Sari: 2013).

Analisa pola frekuensi tinggi dengan algoritma apriori adalah tahap mencari kombinasi *item* yang memenuhi syarat minimum dari nilai *support* sebuah item diperoleh dengan menggunakan rumus berikut: (Kusrini, Emha Taufiq Luthfi, 2010: 150-151).

$$\text{Support (A)} = \frac{\text{Jumlah transaksi mengandung A}}{\text{Total transaksi}} \times 100$$

Untuk mencari nilai *support* dari 2 item diperoleh dengan menggunakan rumus:

$$\text{Support (A,B)} = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi}} \times 100$$

Pembentukan aturan asosiasi adalah aturan setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiasi A -> B. Nilai *confidence* dari aturan A -> B diperoleh dengan rumus berikut: (Kusrini, Emha Taufiq Luthfi, 2010:151).

$$\text{Confidence} = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi mengandung A}} \times 100$$

4. HASIL DAN PEMBAHASAN

Pembahasan dalam penelitian ini adalah dengan menerapkan algoritma apriori pada data transaksi, sehingga data *support* dan *confidence* yang tampil hanya memenuhi persyaratan tersebut.

Analisis Data

Dalam penelitian ini menggunakan sample data transaksi sebanyak 30 transaksi dengan berbagai macam item yang dibeli konsumen. Adapun data tersebut sebagai berikut:

Tabel 1
Daftar Transaksi Toko

No	Tgl	no_trans	nama_barang
1	1 Mei 2017	111	Beras, Gula
2	2 Mei 2017	112	Rokok, Kopi, Telur
3	3 Mei 2017	113	Indomie Rebus, Aqua
4	4 Mei 2017	114	Sabun mandi, Beras, Susu
5	5 Mei 2017	115	Roti, Aqua, Indomie Rebus
6	6 Mei 2017	116	Rokok, Kopi
7	7 Mei 2017	117	Indomie Goreng, Rokok, Gula
8	8 Mei 2017	118	Gula, Gas, Kopi
9	9 Mei 2017	119	Telur, Rokok, Indomie Goreng
10	10 Mei 2017	120	Aqua, Indomie Rebus, Kopi
11	11 Mei 2017	121	Rokok, Telur
12	12 Mei 2017	122	Indomie Rebus, Kopi
13	13 Mei 2017	123	Sabun Mandi, Susu, Roti
14	14 Mei 2017	124	Beras, Indomie Goreng
15	15 Mei 2017	125	Indomie Rebus, Aqua, Rokok
16	16 Mei 2017	126	Telur, Gula, Kopi
17	17 Mei 2017	127	Rokok, Gas, Beras
18	18 Mei 2017	128	Gula, Roti, Indomie Rebus
19	19 Mei 2017	129	Rokok, Indomie Goreng, Susu
20	20 Mei 2017	130	Beras, Gula
21	21 Mei 2017	131	Indomie Goreng, Gas, Kopi
22	22 Mei 2017	132	Telur, Indomie Rebus, Rokok
23	23 Mei 2017	133	Sabun Mandi, Gas, Beras
24	24 Mei 2017	134	Telur, Kopi, Indomie Goreng
25	25 Mei 2017	135	Aqua, Rokok

26	26 Mei 2017	136	Indomie Rebus, Roti, Indomie Goreng
27	27 Mei 2017	137	Beras, Aqua, Telur
28	28 Mei 2017	138	Indomie Goreng, Rokok
29	29 Mei 2017	139	Kopi, Indomie Rebus
30	30 Mei 2017	140	Rokok, Susu, Telur, Beras

Analisis Pola Frekuensi Tinggi

Sebelum dilakukan pencarian pola dari data transaksi terlebih dulu, dicari semua nama jenis item Jenis Barang yang ada didalam transaksi seperti pada **Tabel.1** sekaligus menentukan support per item jenis Jenis Barang, dimana tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai support dalam database :

Berikut merupakan tabel dari semua jenis *itemset* Jenis Barang didalam transaksi penjualan barang (data barang keluar), seperti yang ditunjukkan pada tabel 2 sebagai berikut:

Tabel 2
Daftar Jenis *items* Jenis Barang

No	Nama Item	Support	Support %
1	Aqua	6	7.41
2	Beras	8	9.88
3	Gas	4	4.94
4	Gula	6	7.41
5	Indomie Goreng	8	9.88
6	Indomie Rebus	9	11.11
7	Kopi	9	11.11
8	Rokok	12	14.81
9	Roti	4	4.94
10	Sabun mandi	3	3.7
11	Susu	4	4.94
12	Telur	8	9.88
Jumlah		81	

Data diatas menggambarkan bentuk data 1 *item* yang terdiri atas *atribute item* sebagai nama *item* jenis semua jenis barang yang ada didalam transaksi, support yaitu jumlah setiap item yang ada disemua transaksi, sedangkan *support* (%) yaitu adalah persentasi jumlah *item* yang ada didalam transaksi, yang didapat dari jumlah item dibagi jumlah semua transaksi yang akan dianalisis di kali seratus persen.

Sedangkan Tabel.3 adalah *item* data yang terpilih dengan minimal *support* adalah 9 persen (%). Seperti terlihat pada tabel 3 berikut ini:

Tabel 3
Daftar Jenis *items* dengan *support* yang telah ditentukan

No	Nama Item	Support	Support %
1	Beras	8	9.88
2	Indomie Goreng	8	9.88
3	Indomie Rebus	9	11.11
4	Kopi	9	11.11
5	Rokok	12	14.81
6	Telur	8	9.88

Tabel diatas merupakan data *item* jenis barang yang terseleksi atau terpilih sesuai dengan *support* yang telah ditentukan. Dimana data diatas akan digunakan untuk membentuk pola atau *kombinasi item* dan juga digunakan untuk menentukan *support* dan *confidence* pada pembahasan selanjutnya.

Pembentukan Pola Kombinasi dua items

Pembentukan pola frekuensi dua *item*, dibentuk dari *items-items* jenis barang yang memenuhi *support* minimal yaitu dengan cara mengkombinasi semua *items* kedalam dua kombinasi, hasil dari kombinasi dua *items* seperti pada table 4 berikut ini:

Tabel 4
Daftar Calon Pola Kombinasi Dua *itemset*

No	Nama Item	Support	Support %
1	Beras,Indomie Goreng	1	1.23
2	Beras,Indomie Rebus	0	0
3	Beras, Kopi	0	0
4	Beras, Rokok	2	2.47
5	Beras, Telur	2	2.47
6	Indomie Goreng, Indomie Rebus	1	1.23
7	Indomie Goreng, Kopi	2	2.47
8	Indomie Goreng, Rokok	4	4.94
9	Indomie Goreng, Telur	2	2.47
10	Indomie Rebus, Kopi	3	3.7
11	Indomie Rebus, Rokok	2	2.47
12	Indomie Rebus, Telur	1	1.23
13	Kopi, Rokok	2	2.47
14	Kopi, Telur	3	3.7
15	Rokok, Telur	5	6.17

Data diatas merupakan calon kombinasi dua *item* yang merupakan hasil dari semua kombinasi semua jenis *item*. Dengan menetapkan *support* minimal sama dengan 3% persen, maka data diatas terseleksi atau terpilih, seperti pada tabel 5 berikut ini:

Tabel 5
Daftar Pola kombinasi dua *items* yang memenuhi *support* minimal 3 %

No	Nama Item	Support	Support %
1	Indomie Goreng, Rokok	4	4.94
2	Indomie Rebus, Kopi	3	3.7
3	Kopi, Telur	3	3.7
4	Rokok, Telur	5	6.17

Data diatas adalah kombinasi pola dua *item* data yang terpilih dengan *support* yang terbanyak, itu menandakan bahwa kombinasi dua *item* data tersebut paling banyak didalam transaksi.

Pembentukan Pola Aturan Asosiasi

Setelah semua pola frekuensi tinggi ditemukan, baru dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiasi A ke B.

Proses mencari jumlah kombinasi dan kuatnya hubungan antara satu item dengan item yang lain dalam satu kombinasi disebut metode *association rule*. pembentukan *association rule* adalah menganalisis pola frekuensi tinggi, tahap ini mencari kombinasi yang memenuhi *support* yang telah ditentukan, Terlihat data kombinasi jenis yang memiliki syarat minimum dari *support* dalam database. Pembentukan aturan *association rule*, dengan mencari nilai *confidence*. Dimana *support* adalah jumlah dari kombinasi antara suatu *item* dengan *item* yang lain sedangkan *confidence* adalah nilai yang mendefinisikan kuat tidaknya hubungan antara *item-item* tersebut.

Dari tabel 5 yaitu tabel pola kombinasi dua *item*, dapat dilihat besarnya nilai *support* dan *confidence* dari calon aturan asosiasi seperti tampak pada tabel 6 berikut ini:

Tabel 6
Daftar Calon Aturan Asosiasi dari tabel 5

No	Kombinasi Nama Item	Confidence	Confidence %
1	Jika membeli Indomie Goreng maka akan membeli Rokok	4/8	50
2	Jika membeli Rokok maka akan membeli Indomie Goreng	4/12	33.33
3	Jika membeli Indomie Rebus maka akan membeli Kopi	3/9	33.33

4	Jika membeli Kopi maka akan membeli Indomie Rebus	3/9	33.33
5	Jika membeli Kopi maka akan membeli Telur	3/9	33.33
6	Jika membeli Telur maka akan membeli Kopi	3/8	37.5
7	Jika membeli Rokok maka akan membeli Telur	5/12	41.67
8	Jika membeli Telur maka akan membeli Rokok	5/8	62.5

Dengan menetapkan nilai confidence minimum adalah 40 persen(%), maka aturan yang bisa terbentuk adalah aturan seperti pada tabel 7 berikut ini:

Tabel 7
Daftar Aturan Asosiasi yang terbentuk dari table 6

No	Kombinasi Nama Item	Confidence	Confidence %
1	Jika membeli Indomie Goreng maka akan membeli Rokok	4/8	50
2	Jika membeli Rokok maka akan membeli Telur	5/12	41.67
3	Jika membeli Telur maka akan membeli Rokok	5/8	62.5

5. KESIMPULAN

Berdasarkan pembahasan yang telah dilakukan pengujian dapat diambil beberapa kesimpulan bahwa:

1. Hasil dari pengolahan data 30 transaksi melalui aplikasi *Market Basket Analysis* terdapat 1 pola *asosiasi* yang memenuhi syarat, salah satunya adalah jika membeli telur maka akan membeli rokok dengan nilai *confidence* tertinggi = 62,5%.
2. Pola yang didapat bisa digunakan membantu toko kelontong alief dalam menentukan keputusan memberikan paket diskon atau *bundling* terhadap pola pembelian.

Saran

1. *Data mining* dengan algoritma apriori memiliki kelemahan karena harus melakukan *scan* database setiap kali iterasi, sehingga untuk database yang sangat besar membutuhkan waktu yang lama.
2. Penerapan algoritma apriori sangat praktis namun perlu dilakukan perbandingan dengan algoritma yang lain, untuk menguji sejauh mana

Algoritma Apriori masih dapat diandalkan untuk memproses dan menemukan pola hubungan (*asosiasi*) antar *item* pada database berskala besar.

DAFTAR PUSTAKA

- Dinata, Riadi Marta. (2017). Perbandingan Algoritma Apriori dan Fp-Growth Untuk Rekomendasi Pemasaran dan Tata Letak Pada Transaksi Penjualan Sembako Di Toko “ UcoK” Kampung Sawah. Jurnal Mahasiswa Pasca Sarjana STMIK Nusa Mandiri Kelas 14.1A.01. Jakarta.
- Hidayat, Syamsu. (2016). Penentuan Tren MODE Barang Dengan Algoritma Apriori (Toko Serba ada “Amanah”).
- Kusrini dan Luthfi, Ehma Taufiq. (2010). Algoritma Data Mining, Yogyakarta: Penerbit ANDI Yogyakarta.
- Sari, Eka Novianti. (2013). Analisa Algoritma Apriori Untuk Menentukan Merek Pakaian Yang Paling Diminati Pada Mode Fashion Group Medan.
- Witten, I. H. (2011). *Data Mining Practical Machine Learning Tools and Techniques*. Burlington: Elsevier.