

Pola Pembelian Produk Berdasarkan Association Rule Data Mining

Ratih Wulandari, Ichsani Mursidah
Program Doktor S3 Teknologi Informasi, Universitas Gunadarma
Jl. Margonda Raya No. 100, Depok, Jawa Barat 16424
{ratih_wulandari, mursidah}@staff.gunadarma.ac.id

ABSTRAK

Pengelompokan produk pada penyusunan rak display memiliki pengaruh yang signifikan pada perilaku belanja konsumen. Penelitian ini akan membahas mengenai promosi produk sebagai daya tarik pembelian pelanggan untuk membeli suatu produk menggunakan teknik data mining yaitu Association Rule Mining. Mengidentifikasi pola konsumen dalam membeli barang sangat penting dikarenakan dapat membantu rekomendasi promosi produk sehingga strategi pengelompokan barang pada rak display menjadi lebih tepat. Memanfaatkan data transaksi perhitungan asosiasi produk dilakukan menggunakan software RapidMiner. Hasil dari penelitian ini adalah rule yang dapat digunakan untuk membantu menentukan rekomendasi promosi suatu produk secara lebih tepat, yaitu pola pembelian konsumen hasil asosiasi dengan menggunakan algoritma apriori dengan menggunakan minimum support 90% dan minimum confidence 80% didapat jenis barang yang paling banyak terjual adalah Bumbu Masak --> [Daging] (confidence: 0.857), [Buah_Sayur] --> [Daging] (confidence: 0.857), [Daging] --> [Bumbu Masak] (confidence: 0.960), [Daging] --> [Buah_Sayur] (confidence: 0.960), [Bumbu Masak] --> [Buah_Sayur] (confidence: 0.964) dan [Buah_Sayur] --> [Bumbu Masak] (confidence: 0.964)

Kata Kunci : *Data Mining, Association Rule*

PENDAHULUAN

Persaingan kompetitif antara pengusaha ritel dengan bisnis online yang sedang menjamur membuat para pengusaha ritel harus lebih mengedepankan strategi pemasaran yang tepat guna menarik daya beli konsumen. Salah satu yang harus ditingkatkan adalah terkait promosi produk yang merupakan model aplikasi dari hasil observasi terhadap keadaan dan keinginan pelanggan untuk membeli suatu produk. Rekomendasi Promosi Produk memanfaatkan pendapat pelanggan terhadap suatu barang untuk membantu pelanggan dalam memilih produk. Karena itu Rekomendasi Promosi Produk memerlukan model rekomendasi yang tepat agar apa yang direkomendasikan sesuai dengan keinginan pelanggan, serta mempermudah pelanggan mengambil keputusan yang tepat dalam menentukan produk yang akan dibelinya.

Data mining berguna untuk memberikan solusi kepada para pemilik perusahaan untuk mengambil keputusan untuk meningkatkan bisnis perusahaan. Dengan menggunakan analisis asosiasi dan algoritma apriori dapat menghasilkan suatu rekomendasi yang di harapkan dapat membantu manager dalam mendukung

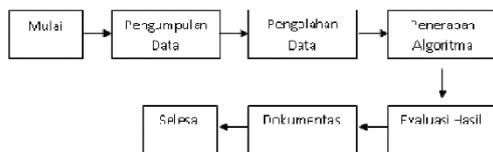
keputusan strategi penjualan. Jadi data mining adalah suatu proses tentang pengambilan data ataupun informasi dalam skala besar dan kemudian dapat diproses menjadi sebuah pengetahuan untuk mendapatkan informasi yang akurat tanpa melalui perhitungan manual [3]. Salah satu hal penting dalam teknik data mining adalah aturan untuk menemukan pola frekuensi tinggi antar himpunan itemset yang disebut dengan fungsi Association Rules (aturan asosiasi). Perkembangan teknologi pada data mining tersebut dapat digunakan untuk menganalisa data dan memprediksi jumlah barang [7]. Analisis keranjang pasar (juga dikenal sebagai *Association Rule Mining*) merupakan salah satu metode data mining yang berfokus pada menemukan pola pembelian dengan mengekstraksi asosiasi atau kejadian dari data transaksional sebuah toko [1]. Selain itu, analisis keranjang pasar adalah cara yang baik untuk memberikan dukungan keputusan ilmiah pada pasar retail melalui hubungan *mining association* antara barang yang telah dibeli secara bersama-sama [2]. Penemuan dari hubungan ini dapat membantu

pedagang untuk mengembangkan strategi penjualan dengan mempertimbangkan barang yang sering dibeli bersama oleh pelanggan [3], hal ini sangat penting karena dapat membantu rekomendasi produk dan promosi produk sehingga strategi pemasaran menjadi lebih tepat. Penentuan pola pembelian barang yang kurang akurat dapat mengakibatkan kebijakan rekomendasi produk dan promosi produk menjadi tidak tepat sasaran.

Banyak algoritma yang diusulkan untuk menemukan pola asosiasi seperti Christidist et al. melakukan penelitian pada *e-commerce* dengan menggunakan *latent topic models* pada data transaksi pasar pelanggan sehingga dapat membantu menentukan rekomendasi produk kepada pelanggan menjadi lebih efektif bahkan pada dataset dan itemset yang besar, Xian Wen et al. menentukan *association rule* dengan menggunakan *matrix database*. Pada penelitian ini akan menggunakan algoritma *FP-Growth* untuk pendekatan asosiasi, sehingga akan ditemukan rekomendasi yang tepat.

METODOLOGI PENELITIAN

Metode penelitian dalam penelitian ini menggunakan metode eksperimen dengan tahapan : (1) Pengumpulan data, (2) Penerapan algoritma, (3) Evaluasi hasil, (4) Dokumentasi eksperimen. Adapun skema alur tahapan penelitian ini ditunjukkan pada gambar 1



Gambar 1. Skema alur tahapan penelitian

2.1. Metode Pengumpulan Data

Metode yang digunakan adalah metode Survei Metode Pengumpulan Data Metode yang digunakan adalah metode Survei dimana mengacu kepada metode pengamanan dimana mengacu kepada metode pengamanan informasi mengenai fenomena yang diteliti dari

seluruh populasi ataupun sampel yang mewakili melalui pengamatan.

2.2. Pengolahan Data

Pengolahan Data yang telah didapat dan dikumpulkan selama proses pencarian dan pengumpulan data kemudian akan masuk ketahap pengerjaan. Metode yaitu dengan proses membedakan data sesuai dengan jenisnya, adapun macam dan jenis data antara lain data primer dan data sekunder. Disini data yang digunakan adalah data Sekunder Perngertian atau penjelasan data sekunder merupakan data yang didapat pencarian data online shop retail yang berbasis di inggris melalui internet, data ini lah yang nantinya akan dijadikan sebagai acuan dalam proses pengembangan sistem data mining. informasi mengenai fenomena yang diteliti dari seluruh populasi ataupun sampel yang mewakili Pengolahan Data yang telah didapat dan pulkan selama proses pencarian dan pengumpulan data kemudian akan masuk ketahap pengerjaan. Metode pengolahan data ini yaitu dengan proses membedakan data sesuai dengan jenisnya, adapun macam dan jenis data antara lain data primer dan data sekunder. Disini data yang digunakan adalah data Sekunder Perngertian atau penjelasan data sekunder merupakan data yang didapat dari proses pencarian data online shop retail yang berbasis di inggris melalui internet, data ini lah yang nantinya akan dijadikan sebagai acuan dalam proses pengembangan sistem data mining.

2.3. Penerapan Algoritma

Didalam teknik association rule ada 2 parameter antara lain adalah confidence dan support. Support yaitu hasil dari nilai yang ditunjang untuk dilihat dengan tingkatan presentasi dalam itemset tersebut dari keseluruhan transaksi yang dilakukan, Penerapan Algoritma Didalam teknik association rule ada 2 meter antara lain adalah confidence dan support. Support yaitu hasil dari nilai yang ditunjang untuk dilihat dengan tingkatan presentasi dalam itemset tersebut dari keseluruhan transaksi yang dilakukan, sedangkan confidence merupakan suatu hasil kepastian yang dapat dilihat kuat tidaknya relasi dengan itemset

tersebut merupakan association rule.[7][6]. Algoritma apriori yaitu salah satu algoritma yang diajukan Agrawal dan Srikant tahun 1994 yang berfungsi untuk menentukan frequent itemset pada asosiasi boolean [7]

2.4. Evaluasi Hasil

Output atau hasil akhir dari hasil akhir dari beberapa tahapan proses data mining agar pola dimengerti dan dapat diterapkan dengan tepat oleh para pelaku usaha ritel ditampilkan bentuk sederhana. Evaluasi yang berjalan pada tahapan ini merupakan memeriksa ulang dan membandingkan data atau informasi yang dirasa berbeda dengan data fakta yang ada dilapangan. proses data mining agar pola informasi dapat dimengerti dan dapat diterapkan dengan tepat oleh para pelaku usaha ritel ditampilkan dalam bentuk sederhana. Evaluasi yang berjalan pada tahapan ini merupakan memeriksa ulang dan membandingkan data atau informasi yang dirasa berbeda dengan data fakta yang ada dilapangan.

2.5. Dokumentasi

Agar setiap hasil dapat dibuktikan kebenarannya maka perlu menampilkan setiap tahapan dan hasil yang diperoleh sebagai dokumentasi yang nantinya dapat dipergunakan untuk evaluasi jika ada penambahan produk atau perubahan strategi pemasaran.

HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Teknik pengumpulan data dilakukan dengan mengambil dataset keranjang pasar. Dalam dataset mendefinisikan transaksi penjualan pada sebuah supermarket. Terdapat 304 atribut yang 303 merupakan item dan salah satunya merupakan ID transaksi penjualan. Terdiri dari 1361 record transaksi. Adapun contoh dari dataset dapat dilihat pada tabel 1.

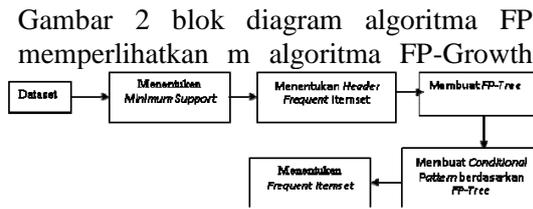
Tabel 1. Tabulasi Data

Data	Beras	Minyak	Margarin	Bumbu	Telur	Susu	Sarapan	Teh	Kopi	Pasta	Buah	Daging	Frozen	Baby	Healthcare	Toiletries	Sanitary	Pe
1	1	1	1	1	1	0	0	0	0	1	1	1	0	0	0	1	0	
2	1	1	0	1	1	0	0	0	0	0	1	0	0	0	1	1	0	
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	
4	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	
5	1	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	
6	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	
7	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	
9	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	
10	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	
11	0	1	0	0	1	1	0	0	0	0	1	0	0	1	0	0	0	
12	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	
15	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	
16	0	1	0	1	1	1	0	0	0	0	0	1	0	0	0	0	0	
17	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
18	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	

Penerapan Algoritma

Algoritma yang biasa dipakai dalam mencari frequent itemset untuk penentuan rekomendasi promosi produk antara lain algoritma dan algoritma FP-Growth. Pada penelitian ini akan dibahas bagaimana pencarian itemset menggunakan algoritma FP-Tree (Frequent Pattern Tree yang biasa dipakai dalam mencari untuk penentuan rekomendasi promosi produk antara lain algoritma Apriori . Pada penelitian ini akan dibahas bagaimana pencarian frequent menggunakan algoritma FP-Growth. Frequent Pattern Tree) digunakan bersamaan dengan algoritma FP-Growth frequent itemset (data yang sebuah dataset. memerlukan langkah , yaitu dengan melakukan bersamaan dengan algoritma untuk menentukan frequent itemset paling sering muncul) dari sebuah Algoritma Apriori memerlukan langkah candidate generation, yaitu dengan melakukan scanning dataset secara berulang menentukan frequent itemset. Algoritma Growth adalah salah satu cara alternatif untuk menemukan himpunan data yang paling sering muncul (frequent itemset) tanpa menggunakan generasi kandidat [5]. secara berulang-ulang untuk . Algoritma FP- adalah salah satu cara alternatif untuk yang paling sering) tanpa menggunakan Untuk menentukan frequent itemset transaksi tersebut, dapat ditunjukkan dengan frequent itemset pada data transaksi tersebut, dapat ditunjukkan dengan bagan berikut:

Gambar 2. Blok Diagram Algoritma Growth



memperlihatkan langkah langkah-langkah dalam menentukan frequent itemset kali dilakukan penentuan dari dataset, setelah itu menentukan frequent itemset dan membuat kemudian FP-Tree yang ada digunakan untuk membuat Conditional Pattern kemudian dilakukan penentuan Frequent Itemset Association Rule Mining matrik: Support, Confidence matrik tersebut berasal dari catatan tr untuk bisnis [6].

a. Support

Matrik pertama ditetapkan untuk analisis keranjang pasar adalah merupakan probabilitas dari asosiasi (probabilitas dari dua item sama).

b. Confidence

Confidence dihasilkan dari seberapa kuat hubungan produk yang sudah dibeli.

c. Lift

Lift Ratio mengukur seberapa penting frequent itemset, dimana pertama kali dilakukan penentuan minimum support dari dataset, setelah itu menentukan header dan membuat FP-Tree, yang ada digunakan untuk Conditional Pattern. Dari kemudian dilakukan Frequent Itemset. Association Rule Mining didasarkan pada tiga Confidence dan Lift.

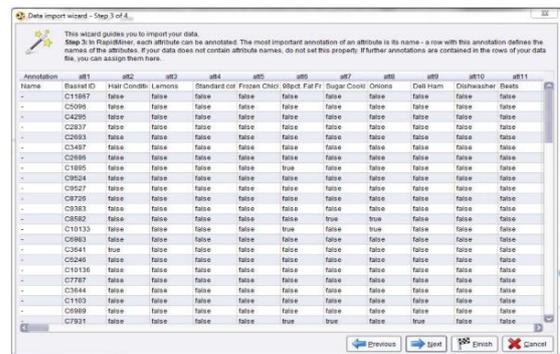
Ketiga matrik tersebut berasal dari catatan transaksi Matrik pertama ditetapkan untuk analisis keranjang pasar adalah Support, yang merupakan probabilitas dari asosiasi item yang dibeli bersama- dihasilkan dari seberapa kuat hubungan produk yang sudah dibeli. mengukur seberapa penting rule yang telah terbentuk berdasarkan nilai support dan telah terbentuk berdasarkan nilai confidence.

Evaluasi Hasil

Hasil keseluruhan dari proses implementasi dalam penentuan rekomendasi produk dengan Algoritma FP-Growth yang disimulasikan pada RapidMiner terhadap data penjualan/transaksi pada dataset memberikan batasan nilai support 40% dan minimum confidence dengan masing-masing nilai Lift > 0,8. Sehingga mendapatkan hasil asosiasi bahwa jenis Hasil keseluruhan dari proses implementasi dalam penentuan rekomendasi produk dengan yang disimulasikan pada RapidMiner terhadap data penjualan/transaksi pada dataset memberikan batasan nilai minimum confidence 80% masing nilai Lift > 0,8. Sehingga mendapatkan hasil asosiasi bahwa jenis barang yang paling banyak terjual adalah “Bumbu Masak, Buah Sayur, Daging, Minyak_Margarin_Mentega, Frozen, Telur dan Beras”. Jadi minimarket ini akan memberikan sebuah informasi/notifikasi kepada supplier yang menjual barang-barang yang paling laku terjual untuk segera menyetok barang tersebut. Asosiasi apriori akan maksimal apabila semakin tinggi nilai minimum support dan minimum confidence maka nilai akurasi semakin tinggi pula, karena hanya item-item yang memiliki asosiasi yang kuat dalam pembelian konsumen yang akan dimunculkan dalam hasil perhitungan.

Dokumentasi Eksperimen

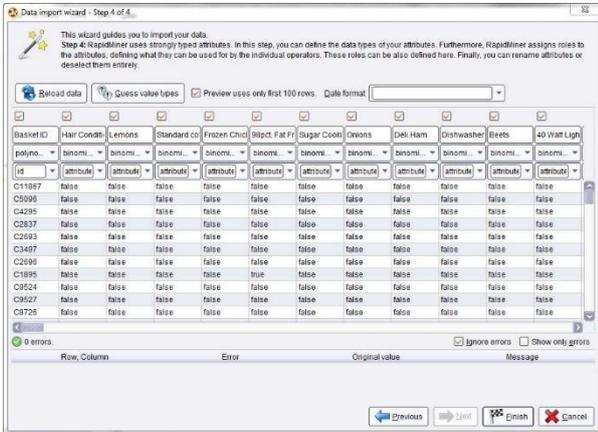
a. Mengambil data dari file .csv yang merupakan dataset analisis keranjang pasar, adapun prosesnya dapat dilihat pada gambar 3.



Gambar 3. Import data dari Dataset

Pada gambar 3 menunjukkan data analisis keranjang pasar yang telah di import kedalam tool rapidminer 5.

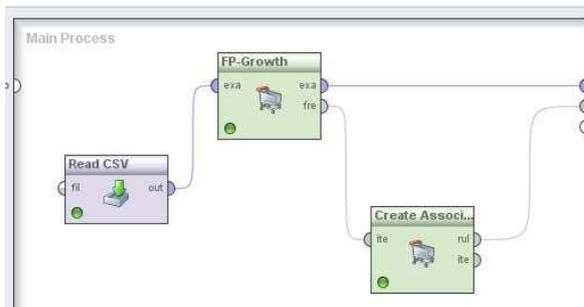
b. Menentukan id



Gambar 4. Menentukan id

Pada gambar 4 memperlihatkan proses pengambilan id sebagai kunci utama (primary key) pada data analisis keranjang pasar.

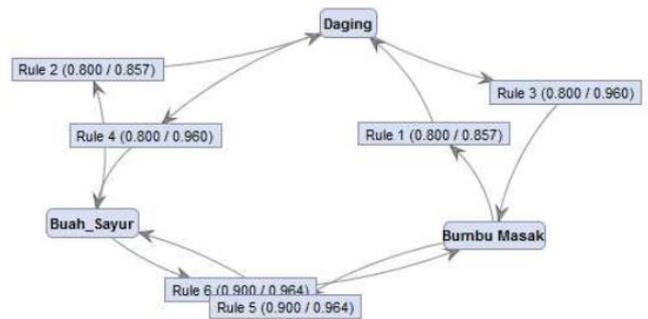
c. Penerapan Algoritma FP-Growth Proses penerapan algoritma dilakukan setelah data diambil dari dataset dan ditentukan id. Gambar 5 menunjukkan penerapan algoritma FPGrowth dalam menentukan rekomendasi promosi produk.



Gambar 5. Penerapan Algoritma FP-Growth

Selanjutnya menentukan nilai minimum support sebesar 90% dan minimum confidence 80% yang paling tepat untuk memperoleh frequent itemset. Pada gambar 5 merupakan proses menggunakan algoritma FP-Growth untuk menentukan rekomendasi promosi produk. Sedangkan gambar 6 menunjukkan hasil pengolahan

data dalam bentuk graphic assosiasi produk menggunakan Rapid Miner.



Gambar 6. Graphic View

KESIMPULAN

Dari hasil penelitian ini dapat disimpulkan bahwa penerapan Association Rule dengan menggunakan Algoritma FP-Growth dapat membantu penentuan rekomendasi promosi produk dengan tepat. Pola pembelian berdasarkan Association Rules yaitu :
[Bumbu Masak]-->[Daging] (conf.: 0.857)
[Buah_Sayur] --> [Daging] (conf.: 0.857)
[Daging] --> [Bumbu Masak] (conf.: 0.960)
[Daging] --> [Buah_Sayur] (conf.: 0.960)
[Bumbu Masak] --> [Buah_Sayur] (conf.: 0.964)
[Buah_Sayur] --> [Bumbu Masak] (conf.: 0.964)

SARAN

Untuk kepentingan lebih lanjut dari penelitian ini, maka penulis memberikan beberapa saran yaitu penelitian selanjutnya dengan metode dan algoritma yang sama peneliti dapat mengembangkannya ke dalam objek yang berbeda, dan dengan menggunakan dataset yang lebih besar. Dengan menggunakan dataset yang lebih besar, peneliti dapat menggunakan nilai support dan confidence yang bervariasi, sehingga dapat memperoleh banyak assosiasi rules antar item.

DAFTAR PUSTAKA

[1]. Agus,Wiwit Triyanto.,2014, "Association Rule Mining Untuk Penentuan Rekomendasi Promosi Produk", Jurnal SIMETRIS, Vol 5, ISSN: 2252-4983.

- [2] . B. Hoanca,. 2011, "Using Market Basket Analysis to Estimate Potential Revenue Increases for a Small University Bookstore", Information Systems, vol. 1822, pp. 1-11.
- [3] . H.K. Kim, J.K. Kim, and Q.Y. Chen. 2012., "A product network analysis for extending the market basket analysis", Expert Systems With Applications, vol. 39, pp. 7403-7410.
- [4] . D.H. Setiabudi, G.S. Budhi, I.W. Purnama, and A. Noertjahyana,. 2011, "Data Mining Market Basket Analysis Using Hybrid-Dimension Association Rules, Case Study in Minimarket X", International Conference on Uncertainty Reasoning and Knowledge Engineering, pp. 196-199.
- [5] . X. Wen-xiu, Q. Heng-nian, and H. Mei-li. 2010., "Market basket analysis based on text segmentation and association rule mining", First International Conference on Networking and Distributed Computing, vol. 309313, pp. 309-313.
- [6] . Lia , Eka Febrianti, Agus Suryadi.,2018, Penerapan Data Mining Dengan Algoritma Apriori Untuk Analisis Pola Pembelian Konsumen, Seminar Nasional Royal (SENAR), STMIK Royal – AMIK Royal, hlm. 173 – 176
- [7] . M. Ykhlef,. 2011, "A Quantum Swarm Evolutionary Algorithm for mining association rules in large databases", Journal of King Saud University - Computer and Information Sciences, vol. 23, pp. 1-6.
- [8] . X. Wu, V. Kumar, J.R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G.J. Mclachlan, A. Ng, B. Liu, P.S. Yu, Z.Z. Michael, S. David, and J.H. Dan., 2008, "Top 10 algorithms in data mining", Knowledge and Information Systems, pp. 1-37.