

Pola Beli Konsumen Menggunakan Algoritma FP-Growth Untuk Rekomendasi Promosi Penjualan Pada Batik Nadya Pekalongan

Edi Faisal, Junta Zeniarja dan Deby Arida NiMatus Sa'adah
Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro
Jl. Nakula I, No. 5-11 Semarang, Kode Pos 50131, Telp. (024) 3515261, Fax: 3569684
edifaisal08@gmail.com, junta@dsn.dinus.ac.id, debyarida07@gmail.com

ABSTRAK

Batik Nadya merupakan toko batik yang berada di Pekalongan. Banyaknya data transaksi pada Batik Nadya yang tersimpan dalam *Microsoft Excel*, menyebabkan penumpukan data. Himpunan data tersebut dapat diolah sehingga akan menghasilkan informasi yang bertujuan meningkatkan volume penjualan melalui strategi rekomendasi promosi produk batik. Data mining adalah ilmu yang dapat menganalisa data dalam jumlah besar sehingga mendapatkan informasi atau pengetahuan. *Association Rule* atau *Market basket analysis* merupakan salah satu teknik dalam data mining untuk menemukan pola pembelian konsumen berdasarkan data transaksi penjualan. Aturan asosiasi memiliki 2 parameter yaitu nilai *support* (nilai penunjang) dan *confidence* (nilai kepastian). Untuk menguji kekuatan aturan asosiasi yang terbentuk adalah dengan mengukur nilai *lift ratio* dari masing-masing rule. Dalam penelitian ini menggunakan Algoritma *FP-Growth* untuk mengolah data transaksi dalam jumlah besar, sehingga menghasilkan *frequent itemsets* melalui pembangunan *FP-Tree* dan menerapkan strategi *divide and conquer*. Dari pengujian yang dilakukan oleh sistem dengan menerapkan *minimum support 2%* dan *min confidence 35%*, menghasilkan tiga aturan asosiasi yang dapat digunakan untuk menentukan rekomendasi promosi produk. Hasil aturan asosiasi yang akan dijadikan target pemberian rekomendasi promosi paket *discount* item berpasangan pada Batik Nadya adalah produk yang memiliki keterkaitan paling tinggi yaitu {hem pria cap, atasan blus wanita} dengan *lift ratio* sebesar 1.74564, *confidence* 0.41042, dan *support* 2.41%.

Kata Kunci : *Association Rule, Market Basket Analysis, FP-Growth, Rekomendasi, Promosi*

1. PENDAHULUAN

Dari persaingan bisnis, khususnya bidang perdagangan menuntut perusahaan atau pengelola bisnis untuk melakukan strategi yang dapat meningkatkan penjualan, sehingga menjamin kelangsungan bisnis tersebut. Batik Nadya merupakan tokoyang menjual berbagai item produk batik secara langsung dan melalui internet, yang berdomisili di Pekalongan. Persaingan yang ketat di Pekalongan dalam bisnis batik, menuntut pegelola Batik Nadya untuk meningkatkan volume transaksi penjualan. Pengelolaan data transaksi direkap pada *Microsoft Excel* yang dari waktu ke waktu data yang tercatat semakin banyak, sehingga terkadang pengelola kesulitan untuk mencari barang apa saja yang sering dibeli konsumen secara bersamaan.

Association Rule atau *market basket analysis* merupakan proses mengidentifikasi hubungan antar suatu item dalam sekumpulan data set yang dibeli oleh pelanggan [1]. Banyak algoritma yang diusulkan untuk menemukan pola asosiasi,

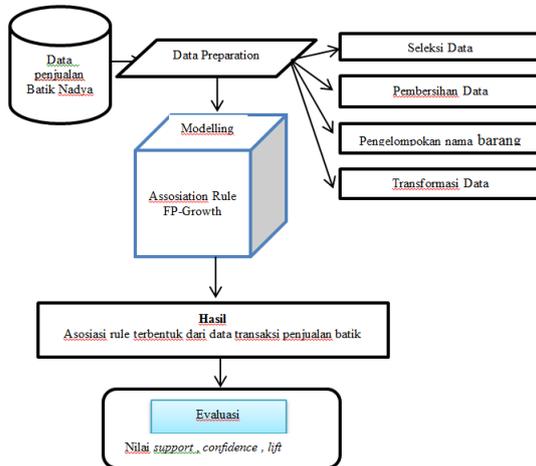
dan algoritma pertama yang efisien adalah algoritma *Apriori*, dengan pemindaian database membutuhkan waktu yang sangat besar[2]. Untuk mengatasi masalah tersebut, algoritma *Frequent Pattern Growth (FP-Growth)* tidak perlu dalam menghasilkan *generate candidate item sets* dan lebih efisien daripada algoritma apriori [3].

FP-Growth merupakan algoritma untuk mencari frequent itemset pada dataset yang besar. Karakter yang dimiliki algoritma *FP-Growth* yaitu menggunakan struktur data tree atau *FP-Tree* sehingga dapat diekstrak secara langsung frequent itemset dari *FP-Tree* tersebut. *FP-Growth* hanya melakukan 2 kali scanning dari data transaksi dan menerapkan strategi *divide and conquer*.

Pada penelitian sebelumnya yang dilakukan oleh Wiwit Agus Triyanto [4] mengenai *association rule mining* untuk penentuan rekomendasi promosi produk, dataset yang digunakan adalah data transaksi penjualan supermarket. Metode yang diusulkan adalah *association rule* dengan menggunakan algoritma *fp-growth*. Hasil yang didapatkan menunjukkan bahwa dalam

Transformasi merubah type data nama kategori produk menjadi angka (nomor kategori produk) dan disimpan dalam file .txt, sehingga dapat diterapkan pada aplikasi.

2.2.4. Modelling



Gambar 3. Modelling

Gambar 3 merupakan *modelling* dalam penelitian, langkah yang pertama data penjualan akan dilakukan processing data, selanjutnya akan diterapkan teknik *modelling* menggunakan *fp-growth*. Menentukan *minimum support* dan *minimum confidence*, kemudian melakukan pembentukan FP-Tree dari data transaksi, membangkitkan *Conditional Pattern Base* yang berisi lintasan prefix dan pola akhiran berdasarkan FP-Tree. Membangkitkan *Conditional FP-Tree* yang memiliki item dengan jumlah *support count* lebih besar dari *Conditional Pattern Base*, dan yang terakhir mencari *frequent itemset*.

2.2.5. Evaluasi

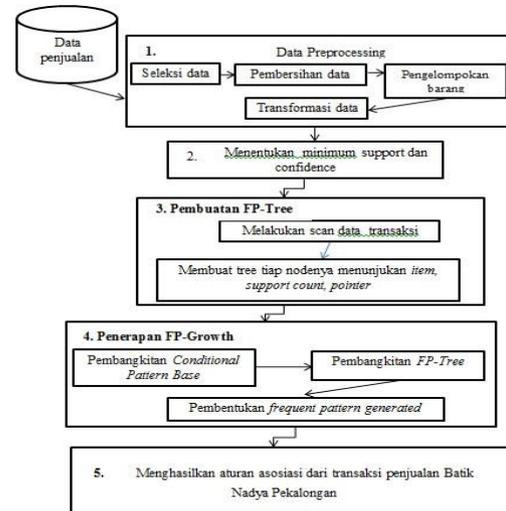
Melakukan pengukuran keakuratan dari hasil yang telah didapatkan. Evaluasi dilakukan untuk menilai pola asosiasi yang dihasilkan dari dataset dengan memperhatikan nilai *support*, *confidence* dan *lift ratio*.

2.2.6. Deployment

Menterjemahkan hasil aturan asosiasi yang didapatkan dari penelitian tersebut, agar dapat dipahami pemilik toko sehingga dapat dijadikan sebagai perencanaan

strategi bisnis dalam memberikan promosi penjualan produk.

2.3 Metode Usulan



Gambar 4. Metode Usulan

1. Data penjualan Batik Nadya akan dilakukan processing data, langkahnya yaitu:
 - a. Seleksi data yang digunakan adalah data dengan transaksi pembelian lebih dari 1 jenis produk.
 - b. Pembersihan data dengan menghilangkan atribut yang tidak digunakan.
 - c. Melakukan pengelompokan barang berdasarkan kategori yang sudah ditentukan.
 - d. Transformasi data yaitu mengubah format data untuk dilakukan proses data mining association rule.
2. Menentukan minimum support dan minimum confidence .
3. Melakukan pembentukan FP-Tree, dengan tahapan berikut :
 - a. Melakukan scan data transaksi penjualan dengan memperhatikan frekuensi item produk.
 - b. Membuat fp-tree berdasarkan data transaksi yang akan diuji.
4. Menerapkan Algoritma FP-Growth, dengan tahapan berikut:
 - a. Membangkitkan Conditional Pattern Base yang berisi lintasan prefix dan pola akhiran berdasarkan FP-Tree.

- b. Membangkitkan Conditional FP-Tree yang memiliki item dengan jumlah support count lebih besar dari Conditional Pattern Base.
 - c. Mencari frequent pattern generated.
5. Menghasilkan aturan asosiasi dari data transaksi penjualan.

3. HASIL DAN PEMBAHASAN

Pengujian evaluasi dilakukan untuk membandingkan frequent itemset yang dibangkitkan algoritma fp-growth dan jumlah rule yang di hasilkan dengan menggunakan minimum support dan minimum confidence yang berbeda pada dataset yang sama sebanyak 5223 transaksi.

3.1 Analisis Pengaruh min sup&min conf terhadap jumlah frequent itemset

Tabel 1.

Pengaruh *min support* dan *min confidence* yang berbeda-beda terhadap *frequent itemset*

Min support	Min Confidence					
	10%	20%	25%	30%	35%	40%
1%	99	99	99	99	99	99
2%	53	53	53	53	53	53
3%	36	36	36	36	36	36
4%	29	29	29	29	29	29
5%	21	21	21	21	21	21
6%	16	16	16	16	16	16
7%	14	14	14	14	14	14

Dari tabel 1, dapat diketahui untuk minimum support terkecil yaitu 1% akan melakukan pembangkitan jumlah frequent itemset terbanyak, sedangkan jumlah frequent itemset paling sedikit akan dibangkitkan minimum support terbesar yaitu 7%. Dalam hal ini, itemset yang memiliki nilai support \geq nilai min support merupakan frequent itemset. Semakin besar min support maka jumlah frequent itemset nya akan semakin sedikit yang dibangkitkan.

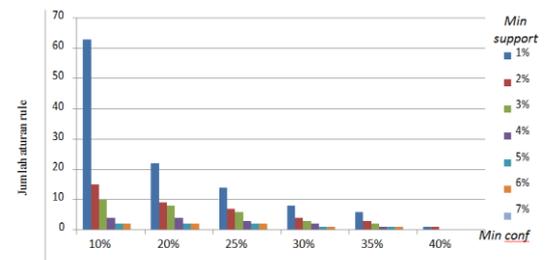
3.2 Analisis Pengaruh min sup &min conf terhadap jumlah aturan asosiasi yang dihasilkan

Tabel 2.

Pengaruh min support dan min confidence yang berbeda-beda terhadap jumlah aturan rule yang dihasilkan

Min support	Min Confidence					
	10%	20%	25%	30%	35%	40%
1%	63	22 aturan	14 aturan	8 aturan	6 aturan	1 aturan
2%	15	9 aturan	7 aturan	4 aturan	3 aturan	1 aturan
3%	10	8 aturan	6 aturan	3 aturan	2 aturan	0 aturan
4%	4	4 aturan	3 aturan	2 aturan	1 aturan	0 aturan
5%	2	2 aturan	2 aturan	1 aturan	1 aturan	0 aturan
6%	2	2 aturan	2 aturan	1 aturan	1 aturan	0 aturan
7%	0	0 aturan	0 aturan	0 aturan	0 aturan	0 aturan

Dari tabel 2, dapat dilihat hasil pengujian untuk masukan minimum support 1%-2% menghasilkan 1 aturan pada minimum confidence 40%. Pada pengujian dengan minimum support 3%-6% ada hal yang menarik yaitu aturan asosiasi tidak dihasilkan pada nilai minimum confidence 40%. Sedangkan pada 7% aturan asosiasi tidak dihasilkan.



Gambar 5. Grafik hubungan jumlah aturan yang dihasilkan dengan nilai *min support* 1%-7% dan *min conf* 10%-40% minimum support 7%

Gambar 5 menunjukkan bahwa nilai min support dan min conf semakin tinggi, akan menyebabkan rule yang terbentuk semakin sedikit. Berdasarkan grafik pengujian, jumlah aturan rule berkurang secara signifikan pada saat min support lebih dari 5% dan min conf lebih dari 25%, rule yang dihasilkan dari minimum tersebut hanya 1 rule, sehingga aturan tersebut tidak ideal. Pada min support 7% tidak menghasilkan rule, karena nilai confidence dari frequent itemset yang dibangkitkan tidak memenuhi syarat minimum conf. Untuk nilai min support yang baik adalah kurang dari 5% dan min conf lebih dari 25%. Kesimpulan yang diperoleh menunjukkan bahwa semakin kecil nilai minimum support maka semakin banyak aturan yang dihasilkan, sebaliknya

semakin besar nilai minimum confidence aturan yang dihasilkan semakin sedikit.

3.3 Analisis nilai lift ratio dari Aturan asosiasi yang dihasilkan

Tabel 3. Nilai lift ratio dari Rule yang dihasilkan

Min supp	Min conf	Aturan asosiasi	Confidence	Lift ratio
2%	30%	{ 1,20 }	0.38095	1.62029
		{ 2,20 }	0.41042	1.74564
		{ 4,20 }	0.31716	1.34898
		{ 24,20 }	0.36752	1.56316
2%	35%	{ 1,20 }	0.38095	1.62029
		{ 2,20 }	0.41042	1.74564
		{ 24,20 }	0.36752	1.56316
2%	40%	{2,20 }	0.41042	1.74564
3%	30%	{ 1,20 }	0.38095	1.62029
		{ 4,20 }	0.31716	1.34898
		{ 24,20 }	0.36752	1.56316
3%	35%	{ 1,20 }	0.38095	1.62029
		{ 24,20 }	0.36752	1.56316
4%	30%	{ 1,20 }	0.38095	1.62029
		{ 4,20 }	0.31716	1.34898
4%	35%	{ 1,20 }	0.38095	1.62029
5%	30%	{ 1,20 }	0.38095	1.62029
5%	35%	{ 1,20 }	0.38095	1.62029
6%	30%	{ 1,20 }	0.38095	1.62029
6%	30%	{ 1,20 }	0.38095	1.62029

Dari pengujian yang sudah dilakukan, untuk penelitian ini *min support* yang digunakan adalah 2% karena rule yang dihasilkan ideal yaitu tidak banyak dan tidak sedikit, *minimum conf* yang digunakan adalah yang tertinggi sebesar 35%, dan batasan minimum tersebut menghasilkan nilai *lift ratio* dan *confidence* paling tinggi diantara batasan minimum yang lainnya. Rumus untuk menghitung nilai confidence dan lift ratio adalah :

$$Confidence = \frac{\text{Jumlah transaksi mengandung X dan Y}}{\text{Jumlah transaksi mengandung X}} \quad (1)$$

$$Lift\ ratio = \frac{Sup(X,Y)}{S(X).S(Y)} \quad (2)$$

Aturan asosiasi yang dihasilkan dari pengujian menggunakan min support 2% dan min conf 35%, dapat diartikan sebagai berikut :

Tabel 4. Hasil Pengujian

No	Aturan	Support	Confidence	Lift ratio
1.	Jika beli item 1 maka juga membeli item 20 (Jika membeli hem pria print maka juga membeli atasan blus wanita)	320 (320/5223 = 6.12%)	0.38095 (38.09%)	1.62029
2.	Jika beli item 2 maka juga membeli item 20 (Jika membeli hem pria cap maka juga membeli atasan blus wanita)	126 (126/5223 = 2.41%)	0.41042 (41.04%)	1.74564
3.	Jika beli item 24 maka juga membeli item 20 (Jika membeli rok batik maka juga membeli atasan blus wanita)	172 (172/5223 = 3.29%)	0.36752 (36.75%)	1.56316

Aturan yang terbentuk memiliki nilai *lift ratio* > 1 maka aturan bersifat kuat dan valid. Pola {hem pria cap, atasan blus wanita} akan dijadikan saran rekomendasi produk karena memiliki keterkaitan paling tinggi, dengan nilai *lift* terbesar yaitu 1.74564.

4. SIMPULAN

Berdasarkan pengujian yang telah dilakukan, kesimpulan dalam penelitian ini adalah :

1. Semakin besar nilai min support dan min confidence yang digunakan, maka jumlah frequent itemset dan aturan asosiasi yang dihasilkan semakin berkurang.
2. Min support yang digunakan adalah 2% dan min confidence 35%, karena menghasilkan rule yang ideal ideal serta nilai lift ratio dan confidence-nya paling tinggi. 3 aturan asosiasi yang terbentuk, yaitu {hem pria print, atasan blus wanita}, {hem pria cap, atasan blus wanita}, dan {rok batik, atasan blus wanita} menghasilkan lift ratio > 1.
3. Aturan asosiasi yang memiliki keterkaitan paling tinggi dengan lift terbesar dapat digunakan sebagai dasar pemberian rekomendasi promosi paket discount yaitu {hem pria cap, atasan blus wanita} menghasilkan nilai lift ratio 1.74564, confidence 41.04%, dan support 2.41%. Hasil tersebut menjelaskan, semakin banyak jenis item yang diteliti, semakin kecil nilai support – nya

5. SARAN

Saran untuk penelitian kedepannya adalah :

1. Penelitian tentang market basket analysis diharapkan dapat lebih spesifik yaitu dengan mengetahui waktu pembelian yang akan diteliti pada hari libur atau hari biasa, sehingga informasi yang dihasilkan dapat memberikan pengetahuan yang efektif.
2. Melakukan analisa terhadap data yang berbeda dengan data transaksi yang lebih bervariasi, yang tidak cenderung memiliki transaksi 2 sampai 3 barang saja, sehingga nilai min supp dan min conf bisa lebih besar.
3. Penelitian ini masih memerlukan pengembangan lebih lanjut dengan menggunakan algoritma lainnya,

sehingga akan menghasilkan komparasi yang baik dan efisien.

DAFTAR PUSTAKA

- [1] Kusrini and E. T. Luthfi, *Algoritma Data Mining*. Yogyakarta: Penerbit ANDI, 2009.
- [2] W. A. Triyanto, V. Suhartono, and H.Himawan, "Analisis Keranjang Pasar Menggunakan K-Medoids dan FP-Growth," *Pseudocode*, vol. 2, no. 1, pp. 129–142, 2014.
- [3] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques Third Edition*, Elsevier, 2012.
- [4] W. A. Triyanto, "ASSOCIATION RULE MINING UNTUK PENENTUAN REKOMENDASI PROMOSI PRODUK," *SIMETRIS*, vol. 5, no. 2, pp. 121–126, 2014.