

Perbandingan Algoritma Apriori Dan Algoritma Fp-Growth Untuk Rekomendasi Item Paket Pada Konten Promosi

Muhammad Mariko¹, Kusri², Sudarmawan³
Universitas Amikom Yogyakarta^{1,2,3}

muhammadmariko@gmail.com, kusri@amikom.ac.id, sudarmawan@amikom.ac.id

Abstrak – Perkembangan teknologi memberi dampak yang begitu besar bagi kehidupan manusia, selain meningkatkan produktivitas teknologi juga dapat memudahkan berbagai pekerjaan termasuk salah satunya yaitu dapat digunakan sebagai sarana yang untuk membantu dalam mengambil sebuah keputusan yang tepat seperti menentukan produk yang tepat untuk digunakan sebagai konten promosi, sehingga konten promosi yang dibuat menjadi lebih tepat sasaran. Pemanfaatan teknologi dalam pengolahan data yang besar bisa menggunakan data mining untuk menganalisa data transaksi penjualan untuk mengetahui pola-pola kombinasi keterhubungan setiap produk sehingga hasil kombinasi tersebut dapat digunakan untuk konten promosi. penelitian ini menggunakan metode apriori dan fp-growth untuk menemukan kombinasi atau keterhubungan setiap produk berdasarkan nilai support dan confidence dari setiap proses mining. Hasil dari penelitian ini menunjukkan bahwa algoritma FP-Growth membutuhkan waktu yang lebih sedikit dibandingkan dengan algoritma Apriori dalam proses mining, kemudian pola asosiasi yang terbentuk dengan nilai minimum support 15% dan nilai minimum confidence 20%. selain itu semakin besar nilai support dan nilai confidence yang diatur maka akan semakin singkat waktu pemrosesan algoritma dan semakin kecil nilai support dan nilai confidence maka akan semakin lama waktu yang dibutuhkan untuk memproses algoritma.

Kata kunci: Data Mining; Association rules; apriori; FP-Growth; konten promosi;

Abstract - Technological developments have a huge impact on human life, in addition to increasing technological productivity it can also facilitate various jobs, including one that can be used as a means to help make the right decisions such as determining the right product to use as promotional content, so that content promotions made to be more targeted. The use of technology in large data processing can use data mining to analyze sales transaction data to find out the combination patterns of the relationship between each product so that the combination results can be used for promotional content. This study uses a priori and fp-growth methods to find the combination or compatibility of each product based on the support and confidence values of each mining process. The results of this study indicate that the FP-Growth algorithm takes less time than the Apriori algorithm in the mining process, then an association pattern is formed with a minimum support value of 15% and a minimum confidence value of 20%. Besides that, the greater the support value and the confidence value that are set, the shorter the algorithm processing time will be and the smaller the supply value and confidence value, the longer it will take to process the algorithm..

Keyword : Data mining; Association rules; apriori; FP-Growth; Promoted Content;

1. Latar Belakang

Promosi merupakan salah satu komponen dari bauran pemasaran (marketing mix). Promosi dapat juga diartikan sebagai upaya untuk memberitahukan atau menawarkan produk atau jasa pada dengan tujuan menarik calon konsumen untuk membeli atau mengkonsumsinya. Dengan adanya promosi produsen atau distributor mengharapkan kenaikannya angka penjualan. Fungsi promosi dalam bauran pemasaran adalah untuk mencapai berbagai tujuan komunikasi dengan konsumen [1]. Setiap perusahaan menggunakan berbagai cara dalam memasarkan produk kepada konsumen salah satunya yaitu membuat konten promosi yang menarik dan juga tepat sasaran.

Dewi Media Lestari adalah perusahaan jasa dan manufaktur yang menyediakan berbagai macam produk paket seminar, merchandise, barang promosi, dll. Umumnya perusahaan menjual produk dalam bentuk paket yang mana setiap paket terdiri berbagai macam item. Setiap paket pembelian terkadang memiliki beberapa item yang hampir sama dengan paket lainnya, dari hal tersebut dapat ditemukan pola pembelian setiap item yang saling berhubungan sehingga dari pola tersebut dimanfaatkan menjadi strategi bisnis dalam mengambil keputusan untuk menentukan konten promosi yang relevan. Dengan mengetahui jenis produk yang sering dibeli maka dapat dibuat sebagai sebuah dasar keputusan untuk menentukan produk apa saja yang cocok untuk dipromosikan kepada pelanggan tersebut. Untuk menemukan pola

transaksi tersebut dapat menggunakan data mining.

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan sebagai ekstraksi informasi yang bermanfaat dan pengetahuan yang terkait berbagai basis data besar. Ada beberapa macam pendekatan yang berbeda yang diklasifikasikan sebagai teknik pencarian informasi/pengetahuan dalam KDD. Ada pendekatan kuantitatif, seperti pendekatan probabilistik seperti logika induktif, pencarian pola, dan analisis pohon keputusan[2]. Salah satu teknik data mining yang dapat digunakan dalam menentukan pola keterhubungan antar item transaksi yaitu menggunakan association rule.

2. Kajian Pustaka

a. Data Mining

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan sebagai ekstraksi informasi yang bermanfaat dan pengetahuan yang terkait berbagai basis data besar[3].

Data Mining juga didefinisikan sebagai proses menemukan pola-pola dalam data. Proses ini otomatis atau seringnya semiotomatis. Pola yang ditemukan harus penuh arti dan pola tersebut memberikan keuntungan, biasanya keuntungan secara ekonomi[4].

b. Asosiation Rule

Association rule mining merupakan salah satu teknik dalam data mining yang berguna untuk menemukan pola asosiasi tersembunyi dalam suatu basis data, pola yang dimaksud disini adalah keterkaitan atau korelasi antara tiap item yang berbeda pada setiap record di dalam basis data. Pola asosiasi yang ditemukan nantinya berupa rule – rule dengan masing – masing nilai bobot asosiasinya, rule yang terbentuk biasa dinotasikan dengan $X \rightarrow Y$ dimana X dan Y disini adalah itemset, bobot asosiasi disini berupa nilai support yang menjelaskan berapa kali sebuah itemset tercatat atau muncul dari sejumlah dataset dan nilai confidence yang menjelaskan seberapa kuat hubungan diantara itemset X dan Y [5]. Metodologi dasar aturan asosiasi terbagi menjadi dua tahap [6], yaitu:

1) Analisis pola frekuensi tinggi

Tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai support dalam database. Nilai support (penunjang) yaitu persentase item atau kombinasi item yang ada pada database. Nilai support sebuah item diperoleh dengan rumus berikut :

$$\text{Support (A)} = \frac{\text{Jumlah transaksi A}}{\text{Jumlah transaksi}}$$

Sedangkan nilai dari support 2 item diperoleh dari rumus berikut :

$$\text{Support (A,B)} = \frac{\text{Jumlah transaksi A dan B}}{\text{Jumlah transaksi}}$$

2) Pembentukan aturan asosiatif

Setelah semua pola frekuensi tinggi ditemukan, maka dicari aturan asosiatif yang memenuhi syarat minimum untuk confidence dengan menghitung confidence aturan asosiatif “Jika A maka B” = $(A \rightarrow B)$.

$$\text{Confidence} = P(B|A) = \frac{\sum \text{Transaksi A dan B}}{\sum \text{Total transaksi A}}$$

c. Apriori

Algoritma apriori adalah langkah untuk proses menemukan frequent-itemset dengan melakukan iterasi pada data. Dimana itemset adalah himpunan item-item yang berada di dalam himpunan yang diolah oleh sistem, sedangkan frequent-itemset menunjukkan itemset yang memiliki frekuensi kemunculan lebih dari nilai minimum yang telah ditentukan (ϕ). Pada iterasi ke-k, semua itemset yang ditemukan yang memiliki k item disebut k-itemset. Setiap iterasi terdiri dari dua tahap yaitu pembangkitan kandidat dan pembangkitan rule[7].

d. FP-Growth

Algoritma Frequent Pattern-Growth (FP-Growth) adalah salah satu cara alternatif untuk menemukan himpunan data yang paling sering muncul tanpa menggunakan generasi kandidat. Frequent Pattern-Growth (FP-Growth) membangun konstruksi data Frequent Pattern Tree (FP-Tree) yang sangat dikompresi, dan mengurangi data asli. Algoritma Frequent Pattern Growth (FP-Growth) melakukan scan database yang sama sebanyak dua kali. Scanning database yang pertama, kita dapat memperoleh frequent 1-item-set, dan scanning database yang kedua, kita dapat memfilter database non-frequent item, selebihnya, Frequent Pattern Tree (FP-Tree) dihasilkan secara bersamaan. Akhirnya, dapat diperoleh

aturan asosiasi dengan menggunakan Frequent Pattern Tree (FP-Tree)[8].

3. Metode Penelitian

Penelitian ini menggunakan tahapan analisis data menurut CRISP-DM (Cross-Industry Standard Process for Data mining) yang terbagi dalam 6 tahapan [9] yaitu :

a. Business Understanding Phase (Fase Pemahaman Bisnis)

Penelitian ini bertujuan untuk mengimplementasikan kemudian membandingkan algoritma apriori dan FP-Growth menggunakan data transaksi penjualan pada perusahaan sehingga hasilnya nanti dapat digunakan sebagai dasar pertimbangan dalam membuat konten promosi.

b. Data Understanding (Fase pemahaman data)

Sumber data yang digunakan adalah data transaksi penjualan pada Perusahaan Dewimedialestari pada bulan Januari - desember 2020.

c. Data Preparation Phase (Fase Persiapan Data)

1) Tahap Pembersihan data

Pada tahap pembersihan data (cleaning data) yaitu membersihkan atau menghilangkan record data yang tidak lengkap (kosong) untuk menghindari missing value sehingga nantinya akan memperoleh hasil yang maksimal. Selain itu juga menghapus data transaksi yang hanya ada 1 item transaksi.

2) Pemilihan Field

Dalam penelitian ini field yang digunakan hanya id transaksi, tanggal transaksi dan nama barang. seperti gambar 1.

nama_item	tgl_order	transaksi
PAYUNG LIPAT 2	2020-01-02	341
BLOCKNOTE A5	2020-01-02	341
TAS FURING	2020-01-06	344
BLOKNOTE A6	2020-01-06	344
PENA	2020-01-06	344
BLOKNOTE A5	2020-01-08	351
TAS KANVAS	2020-01-08	351
PENA BOSJEL	2020-01-08	351
FLASHDISK RANTAI	2020-01-08	351
PENA BOSJEL	2020-01-08	353
TAS FURING	2020-01-08	353
TOTEKAG KANVAS	2020-01-08	353
BLOCKNOTE A5	2020-01-08	353
PAYUNG LIPAT	2020-01-08	354
TAS LIPAT	2020-01-08	354
GANCI PIN	2020-01-08	354
TAS BLACU	2020-01-09	355
BOTOL MINUM	2020-01-09	355
PENA BOSJEL	2020-01-09	355
BLOKNOTE A5	2020-01-11	363
TUMBLER INSERT PAPER	2020-01-11	363
TAS FURING	2020-01-11	365

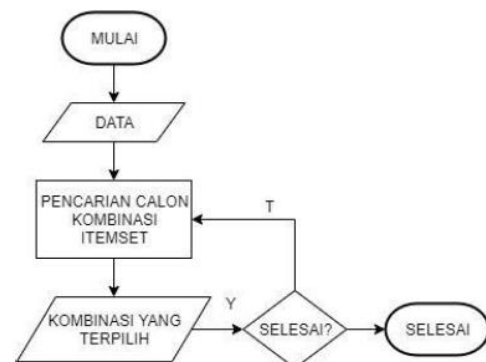
Gambar 1. Data Transaksi Setelah Cleaning

d. Modelling Phase (Fase Pemodelan)

Dalam hal ini dilakukan dengan metode asosiasi data mining menggunakan dua algoritma yaitu apriori dan FP-Growth untuk kemudian dibandingkan.

1) Apriori

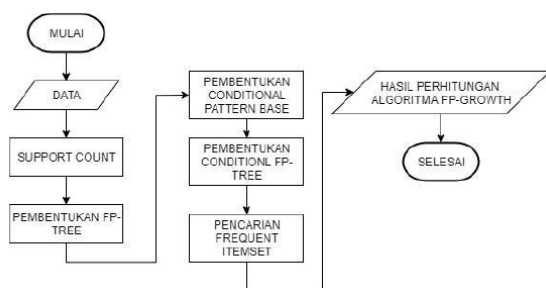
Alur proses dalam algoritma apriori dapat dilihat pada gambar 2.



Gambar 2. Alur Proses Apriori

2) FP-Growth

Alur proses dalam algoritma FP-Growth dapat dilihat pada gambar 3



Gambar 3. Alur Proses FP-Growth

e. Evaluation Phase (Fase Evaluasi)

Tahap ini dilakukan untuk mengevaluasi hasil rule yang terbentuk dari masing-masing algoritma kemudian selanjutnya dilakukan pengambilan kesimpulan dari hasil perbandingan kedua algoritma. Algoritma mana yang paling sesuai untuk menentukan rekomendasi konten promosi pada perusahaan Dewimedialestari.

f. Deployment Phase (Fase Penyebaran)

Pada fase ini, hasil penelitian akan dipresentasikan dalam bentuk laporan yang akan diberikan kepada perusahaan Dewimedialestari untuk dijadikan rekomendasi pemilihan item yang digunakan untuk promosi sehingga dapat membuat konten yang tepat sasaran dan dapat membantu meningkatkan penjualan pada perusahaan Dewimedialestari.

4. Implementasi Sistem dan Hasil

Pada penelitian ini menggunakan uji coba data pada bulan januari 2020. Setelah melewati proses cleaning mendapatkan jumlah transaksi sebanyak 33 paket transaksi dengan jumlah item sebanyak 17 item. minimal jumlah item yang dirposes dalam pembentukan asosiasi rule minimal 2 item dalam satu transaksi. hal ini memudahkan dan mempercepat pemrosesan data hal ini dikarenakan telah dilakukan cleaning terlebih dahulu.

a. Hasil Persentase produk yang sering di beli

Data transaksi akan dialkukan proses perhitungan algoritma untuk mengetahui jumlah transaksi dan support pada setiap item. Setelah itu akan dipilih item yang memenuhi nilai minimum support untuk membuat kombinasi rule 2 item set dan 3 item set. Adapun minimum support yang digunakan adalah sebesar 15% sedangkan untuk minimum confidence nya sebesar 20 %. adapun hasil pemrosesan data tersebut dapat dilihat pada tabel 1.

Tabel 1 . Support 1 item set

No	Item	Jumlah Transaksi	Support Value	%
1	TAS	19	0.58	58
2	PENA	18	0.55	55
3	BLOCKNOTE	16	0.48	48
4	POUCH	9	0.27	27
5	TUMBLER	7	0.21	21
6	FLASHDISK	7	0.21	21
7	PAYUNG	5	0.15	15
8	LANYARD	4	0.12	12
9	TOTEBAG	3	0.09	9
10	BOX	2	0.06	6
11	KAOS	2	0.06	6
12	JAM	2	0.06	6
13	MUG	2	0.06	6
14	LEAFLET	1	0.03	3
15	POSTER	1	0.03	3
16	KEMEJA	1	0.03	3
17	GANCI	1	0.03	3

b. Hasil rule yang diperoleh

Dengan minimum support yaitu 15% atau dan minimum confidence 20% mendapatkan hasil :

- 1) Jika Membeli TAS dan PENA maka pembeli juga membeli BLOCKNOTE dengan confidence 37%
- 2) Jika Membeli PENA dan BLOCKNOTE maka pembeli juga membeli FLASHDISK dengan confidence 28%

5. Kesimpulan

Berdasarkan pembahasan penelitian yang penulis lakukan, dapat disimpulkan bahwa :

- a. Dalam pembentukan rule algoritma fp growth lebih cepat dibandingkan dengan algoritma apriori.
- b. Pola asosiasi yang terbentuk dengan nilai minimum support 15% dan nilai minimum confidence 20% menghasilkan 2 aturan asosiasi. Dan strong rules yang didapatkan adalah "If TAS and PENA then BLOCKNOTE" dengan nilai support 7% dan nilai confidence 37%.
- c. Semakin besar jumlah data yang di proses maka semakin lama waktu yang dibutuhkan untuk membuat rule.
- d. Semakin besar nilai support dan nilai confidence yang diatur maka akan semakin singkat waktu pemrosesan algoritma dan

semakin kecil nilai support dan nilai confidence maka akan semakin lama waktu yang dibutuhkan untuk memproses algoritma.

- e. Analisis pola yang dihasilkan aplikasi ini dapat digunakan oleh perusahaan Dewimedialestari dalam untuk dijadikan rekomendasi pemilihan item yang digunakan untuk membuat konten promosi berdasarkan paket item set.

Rule Dengan Algoritma Fp-Growth Untuk Mengetahui Kelulusan Tepat Waktu Mahasiswa. Jurnal Edik Informatika, V5. i1(41-53) .

- [9] Fadillah, A.P., (2015), "Penerapan Metode CRISP-DM untuk Prediksi kelulusan Studi Mahasiswa Menempuh Mata Kuliah", vol.1, Bandung, Desember, 260-270.

6. Pustaka

- [1] Penda Sudarto Hasugian. 2018. "Perancangan Website Sebagai Media Promosi Dan Informasi". Journal Of Informatic Pelita Nusantara, Vol. 3, no. 1
- [2] Bobby Septia Pranata, Dito Putro Utomo.(2020). "Penerapan Data Mining Algoritma FP-Growth Untuk Persediaan Sparepart Pada Bengkel Motor". Bulletin of Information Technology (BIT), Vol. 1, No. 2, Page 83 - 91.
- [3] Kezia Sumangkut, Arie Lumenta, Virginia Tulenan (2016). "Analisa Pola Belanja Swalayan Daily Mart Untuk Menentukan Tata Letak Barang Menggunakan Algoritma FP-Growth", E-Journal Teknik Informatika Vol. 8, No. 1.
- [4] Haroe Snatoso, I Putu Hariyadi, Prayitno . (2016). "Data Mining Analisa Pola Pembelian Produk dengan Menggunakan Metode Algoritma Apriori." Yogyakarta: Teknik Informatika
- [5] Ulya Anisatur Rosyidah, Hardian Oktavianto. (2018). "Pencarian Pola Asosiasi Keluhan Pasien Menggunakan Teknik Association Rule Mining." Informatics Journal, Vol. 3 No.1.
- [6] Kusriani dan Luthfi, E.T. (2009). Algoritma Data Mining. Yogyakarta : Andi.
- [7] Mohamad Fauzy, Kemas Rahmat Saleh, Ibnu Asror. (2016). "Penerapan Metode Association Rule menggunakan Algoritma Apriori Pada Simulasi Prediksi Hujan Wilayah Kota Bandung". Jurnal Ilmiah Teknologi Informasi Terapan (JITTER), Vol. II, No. 2
- [8] Satrio Junaidi, Thomson Mary. (2018). Metode Data Mining Association