

Implementasi Algoritma Frequent Pattern Growth (FP-Growth) untuk Rekomendasi Pembelian Barang

(Studi Kasus : PD. Indoapi Sarana Utama)

Arief Sasono¹, Andika Elok Amalia², Rima Dias Ramadhani³

^{1,3} Program Studi S1 Informatika, Fakultas Teknologi Industri dan Informatika, Institut Teknologi Telkom Purwokerto

² Program Studi S1 Rekayasa Perangkat Lunak, Fakultas Teknologi Industri dan Informatika, Institut Teknologi Telkom Purwokerto

^{1,2,3} Jl. D. I. Panjaitan No. 128, Purwokerto Kidul, Purwokerto Selatan, Banyumas, Jawa Tengah Kabupaten Banyumas, Jawa Tengah 53147

Email : andika.amalia@ittelkom-pwt.ac.id

Abstrak – Pencarian pola asosiasi barang khususnya pada transaksi penjualan maupun ritel saat ini sering digunakan. Metode yang digunakan untuk menganalisis keranjang belanja data transaksi penjualan produk termasuk dalam *data mining* sebagai suatu teknik analisis data yang dapat membantu pihak manajemen memperoleh pengetahuan berupa pola – pola penjualan. Salah satu metode *data mining* yang digunakan dalam penelitian ini adalah *association rule* dengan menerapkan Algoritma Frequent Pattern Growth (FP-Growth). Metode ini dimulai dengan mencari sejumlah frequent itemset dan dilanjutkan dengan pembentukan aturan – aturan asosiasi (*association rules*). Algoritma FP – Growth adalah algoritma yang sangat populer untuk menemukan sejumlah frequent itemset dari data – data transaksi . Penelitian ini menghasilkan 16 aturan asosiasi yang valid dari total 35 aturan asosiasi yang dihasilkan, dengan nilai *support* $\geq 4.43\%$, *confidence* $\geq 35.21\%$ dan *lift ratio* ≥ 1.06 . Aturan-aturan tersebut kemudian dapat dijadikan acuan untuk rekomendasi barang pembelian.

Kata kunci – analisis keranjang belanja, fp-grwoth, apriori, asosiasi, *data mining*

Abstract— Pattern finding on item association, especially in sales and retail transactions, is now often used. The method used to analyze the product sales transaction data shopping cart is included in data mining as a data analysis technique that can help management gain knowledge in the form of sales patterns. One of the data mining methods used by the Frequent Pattern Growth Algorithm (FP-Growth). This method starts with searching for a number of frequent itemsets and continues with the formation of association rules. FP - Growth The algorithm is a very popular algorithm to find a number of frequent items from transaction data. This study produced 16 valid association rules from a total of 35 association rules produced, with the support value $\geq 4.43\%$, confidence $\geq 35.21\%$ and elevator ratio ratio 1.06. These rules can be used as a reference for the purchase of items

Keywords- *market basket analysis, fp-grwoth, apriori, association, data mining*

I. PENDAHULUAN

Market Basket Analysis merupakan sebuah metode untuk menentukan pola hubungan antara barang satu dengan yang lain dan perilaku konsumen dalam transaksi jual beli [1]. Metode tersebut menganalisis hubungan dari produk-produk yang dibeli secara bersamaan atau yang disebut dengan *rule*, serta dapat memberikan saran kepada manajemen untuk persediaan barang dengan melibatkan penggunaan data transaksional konsumen. Perusahaan Dagang (PD) Indoapi Sarana Utama merupakan perusahaan dagang yang bergerak dalam bidang keamanan kantor dan perusahaan. Perusahaan tersebut menjual APAR atau dengan kata lain Alat Pemadam Api Ringan terdiri dari APAR Padat, APAR Cair, APAR Gas,

Alarm Rumah, Alarm Kantor, Brankas Masterindo, Brankas Chubb, Cashbox Eagle, Cashbox Masterindo, CCTV, Helm Proyek, Lemari Arsip Brother, Lemari Arsip Masterindo, Money Detector, Suku Cadang Brankas, Suku Cadang APAR, Hydrant, Pintu Khasanah Bank Chubb, Pintu Khasanah Bank Masterindo, dan Mesin Hitung Uang yang berlokasi di Purwokerto. Sebagai sebuah perusahaan, PD Indoapi Sarana Utama memiliki banyak sekali data transaksi dalam rentang waktu lima tahun (2010-2015) yang dapat dimanfaatkan salah satunya untuk mengetahui pola asosiasi barang.

Penentuan asosiasi barang sudah banyak dilakukan sebelumnya terutama untuk kasus-kasus dalam dunia ritel. Penelitian yang dilakukan oleh Fitri Nurchalifatun yaitu pencarian pola kombinasi antar

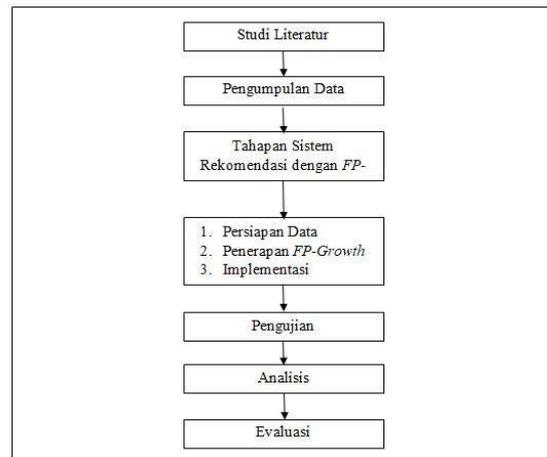
itemset pada Pondok Kopi dengan menerapkan Algoritma Apriori. Hasil dari penelitian tersebut adalah 7 aturan asosiasi, namun masih terdapat kekurangan yaitu jumlah data serta penggunaan confidence dan support yang belum bervariasi sehingga perlu ditingkatkan agar diperoleh hasil aturan asosiasi yang efisien [2]. Penelitian yang juga menggunakan Algoritma Apriori dilakukan oleh D. Anggraeni, H. R. Saputra dan B. Noranita yaitu menganalisis pola asosiasi dalam transaksi penjualan obat. Nilai *minimum support* ditentukan 1,7% dan *minimum confidence* 10% yang menghasilkan nilai *lift ratio* 0,625.

Algoritma lain yang dapat digunakan dalam metode market basket analysis adalah Algoritma Frequent Pattern Growth (FP-Growth). Penelitian yang dilakukan oleh Ikhwan, D.A, Nofriansyah dan Sriani menerapkan algoritma tersebut untuk mendukung promosi pendidikan. Tujuannya adalah untuk menganalisa permasalahan promosi pendidikan di STMIK Triguna Dharma serta menerapkan algoritma FP – Growth dalam strategi promosi pendidikan dengan *tools* yang digunakan Rapidminer untuk menghasilkan *rules* asosiasi [4]. Penelitian lain dengan menggunakan metode Fold-Growth dilakukan oleh A. B. N. Sari, Rahman and Y. T. Mursityo untuk mengetahui pola asosiasi antara data mahasiswa dan tingkat kelulusan dengan nilai *minimum support* 20% dan *minimum confidence* 50% yang menghasilkan nilai *lift ratio* 0,531 [5]. Pada penelitian yang dilakukan oleh A. R. M dilakukan komparasi antara Algoritma Apriori dan FP-Growth. Perbedaan kedua algoritma tersebut terletak pada waktu eksekusi program dimana dengan menggunakan Apriori waktu yang digunakan untuk eksekusi program lebih banyak dibandingkan waktu eksekusi program menggunakan FP - Growth. Dari hasil kinerja kedua algoritma semakin besar minimum supportnya maka *itemsets* dan *Association Rule* yang terbentuk tidak ada, maka dari itu dicari minimum support yang paling mendekati agar frequent itemsets dan Association Rule yang terbentuk sempurna atau mendekati sempurna, dan jumlah data buku yang dipinjam ditingkatkan agar pola mining yang terbentuk semakin bervariasi [6].

Pada penelitian ini digunakan Algoritma FP-Growth karena dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data. Algoritma FP-Growth merupakan pengembangan dari algoritma Apriori, sehingga kekurangan dari algoritma Apriori diperbaiki oleh algoritma FP-Growth. Pada Algoritma FP-Growth *generate candidate* tidak dilakukan karena menggunakan konsep pembangunan *tree* dalam pencarian *frequent itemsets* [7]. Hal tersebut yang menyebabkan Algoritma FP-Growth lebih cepat dari Algoritma Apriori.

II. METODE PENELITIAN

Tahapan penelitian yang akan dilakukan dapat dilihat pada Gambar 1.



Gambar 1. Tahapan penelitian

A. Studi Literatur

Studi Literatur dilakukan untuk mempelajari terkait metode market basket analysis dan algoritma yang dapat diimplementasikan di dalamnya termasuk Algoritma FP-Growth dari penelitian yang pernah dilakukan. Beberapa penelitian yang dijadikan rujukan sudah dibahas pada bab sebelumnya.

B. Pengumpulan Data

Pada tahap ini dikumpulkan data penjualan sejumlah 2500 transaksi pada tahun 2010 hingga 2015 yang akan diolah pada Penerapan Algoritma FP-Growth yang digunakan untuk menemukan frequent itemset dimana nantinya akan menjadi rule atau aturan rekomendasi pembelian barang. Data transaksi tersebut bersumber dari PD Indoapi Sarana Utama.

C. Rekomendasi Menggunakan Algoritma FP-Growth

Tahapan ini menjadi 3 langkah sebagai berikut.

1. Persiapan data yang akan diolah pada tahap dataset yang terdiri dari pembatasan dengan support count hingga pembentukan FP-Tree.
2. Setelah FP-Tree dibentuk, kemudian tahap FP-Growth yang dibagi menjadi 3 tahap, yaitu conditional pattern base, conditional FP-Tree, dan tahap penemuan frequent itemset. Jika *frequent itemset* telah ditemukan maka akan masuk pada tahap *association rule*, pada tahap tersebut frequent itemset yang telah ditemukan akan menjadi rule yang kemudian ditentukan nilai support dan confidence masing-masing.
3. Hasil rule yang ditemukan akan diuji terlebih dahulu, kemudian akan menjadi rekomendasi pembelian yang diterapkan kembali pada sistem penjualan PD Indoapi Sarana Utama.

D. Pengujian

Pengujian yang dilakukan pada penelitian ini menggunakan *lift ratio*, yaitu dengan menghitung nilai *lift ratio* masing-masing *rule* yang telah ditemukan *support* dan *confidence*.

E. Analisis

Hasil *lift ratio* setiap *rule* yang ditemukan pada tahap pengujian akan dianalisis apakah *rule* tersebut valid/kuat untuk dijadikan rekomendasi pembelian barang dan strategi pemasaran di PD Indoapi Sarana Utama

F. Penarikan Kesimpulan

Hasil *rule* yang telah memiliki *lift ratio* yang valid, dilakukan penarikan kesimpulan manakah *rule* yang memiliki nilai *support*, *confidence* dan *lift ratio* yang paling akurat.

III. HASIL PENELITIAN

Pada bab ini akan dibahas proses implementasi FP-Growth hingga menghasilkan *association rules* dari keseluruhan data yang diolah. Formula 1 dan 2 adalah yang digunakan dalam prosesnya.

$$Support(A) = \frac{Jumlah\ Transaksi\ Mengandung\ A}{Total\ Transaksi} \tag{1}$$

$$Support(A,B) = P(A,B) = \frac{Jumlah\ Transaksi\ Mengandung\ A\ dan\ B}{Total\ Transaksi} \tag{2}$$

A. Pembentukan FP-Tree

Dataset atau data transaksi awal dibatasi dengan *support count* yang telah ditentukan, kemudian dibangun menjadi sebuah tree. Berikut langkah-langkah perhitungan untuk mendapatkan nilai *support* pada setiap *item*.

1. Hitung jumlah total setiap item atau barang berdasarkan data transaksi. Setelah jumlah total setiap item diketahui, maka dibagi dengan jumlah total data transaksi yang berjumlah 2500 data, jika hasil telah diketahui maka dikalikan angka 100 untuk mendapatkan hasil prosentasenya (menggunakan Formula 1).
2. Berikut adalah contoh perhitungan untuk mendapatkan nilai *support* pada item “APAR Padat”.
3. Berdasarkan data transaksi jumlah total item “APAR Padat” berjumlah 1080, kemudian 1080 dibagi dengan nilai jumlah total data transaksi yaitu bernilai 2500. Sehingga, $1080 \div 2500 = 0,432$.
4. Angka 0,432 yang merupakan hasil perhitungan untuk mendapatkan nilai *support* dikalikan dengan angka 100 untuk mendapatkan nilai berbentuk prosentase sehingga memiliki nilai *support* bernilai 43.2%.

Keseluruhan hasil perhitungan masing-masing barang dapat dilihat pada Tabel 1.

Tabel 1. Hasil Perhitungan Nilai Support Seluruh Item

Barang	Jumlah	Support	
APAR Padat	1080	$1080 \div 2500$	43.20%
Brankas	988	$988 \div 2500$	39.52%

Barang	Jumlah	Support	
Masterindo			
Suku Cadang APAR	752	$752 \div 2500$	30.08%
Cash Box Eagle	685	$685 \div 2500$	27.40%
Mesin Hitung Uang	671	$670 \div 2500$	26.84%
Hydrant	208	$208 \div 2500$	8.32%
Brankas Chubb	192	$192 \div 2500$	7.68%
Alarm Kantor	186	$186 \div 2500$	7.44%
Suku Cadang Brankas	161	$161 \div 2500$	6.44%
Lemari Arsip Brother	160	$160 \div 2500$	6.40%
Pintu Khasanah Masterindo	150	$150 \div 2500$	6%
Lemari Arsip Masterindo	113	$113 \div 2500$	4.52%
Alarm Rumah	103	$103 \div 2500$	4.12%
CCTV1	95	$95 \div 2500$	4.80%
CCTV2	80	$80 \div 2500$	4.20%
Money Detector	78	$78 \div 2500$	4.12%
Cash Box Masterindo	65	$65 \div 2500$	2.60%
Pintu KBC	64	$64 \div 2500$	2.56%
CCTV3	63	$63 \div 2500$	2.52%
APAR Cair	48	$48 \div 2500$	1.92%
APAR Gas	32	$32 \div 2500$	1.28%
Helm Proyek	21	$21 \div 2500$	0.84%

Penentuan minimum *support count* berdasarkan penelitian terdahulu dan sesuai dengan jumlah data transaksi yang akan digunakan dalam penelitian ini yaitu bernilai 25% [5], sehingga didapatkan item yang akan diolah menjadi *rule* asosiasi adalah yang jumlah keberadaannya dalam transaksi lebih besar dari 625 ($25\% \times 2500$ transaksi). Yaitu APAR Padat (AP),

Brankas Masterindo (BM), Cash Box Eagle (CE), Suku Cadang APAR (SCA) dan Mesin Hitung Uang (MHU).

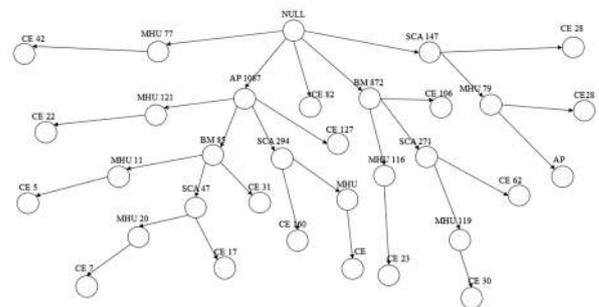
Tabel 2 merupakan data transaksi yang didapatkan dari PD Indoapi Sarana Utama.

Tabel 2. Data Transaksi

No	Transaksi	Jumlah
1	AP	429
2	BM	379
3	AP, CE	127
4	BM, CE	106
5	AP, MHU	100
6	AP, SCA	99
7	BM, MHU	93
8	BM, SCA	90
9	BM, SCA, MHU	89
10	CE	82
11	SCA, MHU, AP	78
12	BM, SCA, CE	62
13	AP, SCA, CE	60
14	SCA, MHU	51
15	MHU, CE	42
16	SCA	40
17	MHU	35
18	AP, BM, CE	31
19	BM, SCA, MHU, CE	30
20	SCA, MHU, CE	28
21	SCA, CE	28
22	AP, BM	27
23	MHU, CE, BM	23
24	MHU, CE, AP	21
25	AP, BM, SCA, CE	17
26	SCA, MHU, CE, AP	16
27	AP, BM, SCA, MHU	13
28	AP, BM, SCA	10
29	AP, BM, SCA,	7

No	Transaksi	Jumlah
	MHU, CE	
30	AP, BM, MHU	6
31	MHU, CE, AP, BM	5
	JUMLAH	2258

Dari Tabel 2 akan dipilih menjadi beberapa untuk diseleksi menjadi *frequent itemset* dimana yang terpilih adalah yang memuat barang-barang hasil tahap sebelumnya. Karena semua transaksi memuat AP, BM, CE, SCA atau MHU maka keseluruhannya dibuat menjadi FP- Tree pada Gambar 2.



Gambar 2. Hasil FP-Tree

B. Pembuatan FP-Growth

Proses penentuan *frequent itemset* mengacu pada barang-barang yang memenuhi minimum support count dapat dilihat pada Tabel 3.

Tabel 3. Frequent Itemset

Suffix	Frequent Itemset
CE	{{CE}, {CE,MHU}, {CE,SCA}, {CE,BM}, {CE,SCA,BM}, {CE,AP}, {CE,MHU,SCA}, {CE,SCA,AP}}
MHU	{{MHU}, {MHU,AP}, {MHU,BM}, {MHU,SCA}}
SCA	{{SCA}, {SCA,BM}, {SCA,AP}}
BM	{{BM}, {BM,AP}}
AP	{{AP}}

C. Hasil Association Rules

Tabel 4 merupakan daftar hasil aturan asosiasi setelah mengolah data pada Tabel 3. Masing-masing aturan asosiasi telah dihitung nilai *support*, *confidence* dan *lift ratio*. Penghitungan lift ratio menggunakan Formula 3.

$$Lift\ Ratio = \frac{Confidence}{Benchmark} \tag{3}$$

Tabel 4. Hasil Association Rules

No	Rule	Support	Confidence	Lift ratio
1	AP, MHU => SCA	6.55%	52.86%	1.59
2	BM, MHU => SCA	6.16%	52.26%	1.57
3	MHU => SCA	15.32%	51.56%	1.55
4	SCA, CE => BM	5.14%	46.77%	1.07
5	AP, SCA => MHU	6.55%	44.31%	1.49
6	BM, SCA => MHU	6.16%	44.71%	1.47
7	BM, CE => SCA	5.14%	41.28%	1.24
8	SCA => MHU	15.32%	46.01%	1.55
9	MHU => AP, SCA	6.55%	22.06%	1.49
10	MHU => BM, SCA	6.16%	20.72%	1.47
11	CE, MHU => SCA	4.59%	47.09%	1.41
12	BM, SCA => CE	5.14%	36.48%	1.2
13	SCA, CE => MHU	4.59%	32.66%	1.1
14	SCA => CE	10.98%	32.98%	1.09
15	CE => SCA	10.98%	36.20%	1.09
16	AP, CE => SCA	4.43%	35.21%	1.06

No	Rule	Support	Confidence	Lift ratio
17	AP, SCA => CE	4.43%	29.94%	0.99
18	BM => SCA	14.08%	32.19%	0.97
19	SCA => BM	14.08%	42.29%	0.97
20	BM => CE	12.44%	28.44%	0.94
21	CE => BM	12.44%	41.02%	0.94
22	AP => SCA	14.79%	30.93%	0.93
23	SCA => AP	14.79%	44.41%	0.93
24	SCA, MHU => BM	6.16%	40.17%	0.92
25	BM => MHU	11.78%	26.92%	0.91
26	MHU => BM	11.78%	39.64%	0.91
27	SCA, MHU => AP	6.55%	42.77%	0.89
28	AP => MHU	12.40%	25.93%	0.87
29	MHU => AP	12.40%	41.73%	0.87
30	AP => CE	12.58%	26.30%	0.87
31	CE => AP	12.58%	41.46%	0.87
32	CE => MHU	7.62%	25.11%	0.84
33	MHU => CE	7.62%	25.63%	0.84
34	SCA, CE	4.43%	40.32%	0.84

No	Rule	Support	Confidence	Lift ratio
	=> AP			
35	SCA, MHU => CE	4.59%	24.41%	0.77

IV. PEMBAHASAN

Berdasarkan Tabel 4, terdapat sejumlah 16 aturan yang memiliki nilai lift ratio lebih besar dari 1 (lift ratio >1). Hal ini menunjukkan bahwa 16 aturan tersebut bersifat kuat dan valid untuk digunakan sebagai acuan dalam merekomendasikan pembelian barang di PD Indoapi Sarana Utama.

Dapat diketahui proses untuk mendapatkan aturan yang valid atau memiliki nilai lift ratio > 1 yaitu dengan nilai-nilai yang terdapat pada Tabel 4. Untuk mendapatkan nilai lift ratio terlebih dahulu harus mengetahui nilai benchmark dari setiap rule. Nilai benchmark dapat dicari dengan membandingkan nilai frekuensi barang consequent dengan total data transaksi yang ada dalam database seperti yang terdapat pada Formula 4.

$$Benchmark\ Confidence = \frac{Nc}{N} \tag{4}$$

Dalam menemukan nilai benchmark dibutuhkan nilai frekuensi barang consequent, nilai tersebut didapatkan dari barang yang menjadi conclusion pada association rule. Kemudian jika nilai benchmark telah diketahui, maka nilai lift ratio dapat dicari dengan membandingkan nilai confidence dengan nilai benchmark confidence seperti yang dijelaskan pada rumus. Sehingga, dapat disimpulkan berdasarkan penelitian serupa sebelumnya dan berdasarkan hasil analisis bahwa penelitian ini telah sukses atau lebih baik dari penelitian-penelitian serupa sebelumnya, karena rule yang dihasilkan berjumlah 16 rule dengan nilai lift ratio > 1, dan pada penelitian-penelitian sebelumnya terbentuk maksimal 7 rule dengan terdapat lift ratio < 1.

V. PENUTUP

A. Kesimpulan

1. Penerapan metode asosiasi dan algoritma FP – Growth telah memberikan output Association Rule valid sejumlah 16 dengan minimum support (minsupp) 0.03 dan minimum confidence (mincof) 0.2.
2. Association Rule awal yang dihasilkan yaitu berjumlah 35 rule dan diuji berdasarkan lift ratio yang bernilai > 1 terdapat 16 rule yang valid dan dapat dijadikan sebuah pegangan dalam strategi penjualan barang serta rekomendasi untuk pembelian barang di PD Indoapi Sarana Utama.
3. Hasil Association Rule yang diperoleh yang memiliki nilai confidence tertinggi adalah 52.86% dengan spesifikasi rule APAR (Alat Pemadam Api Ringan), Mesin Hitung Uang => Suku Cadang APAR.

B. Saran

1. Penelitian ini dapat dikembangkan dengan menggunakan algoritma Association Rule yang lain untuk dilakukan perbandingan.
2. Jumlah data transaksi ditingkatkan agar pola mining yang terbentuk semakin bervariasi.
3. Sistem pemrosesan algoritma FP – Growth dan Association Rule dapat disatukan dalam sistem informasi penjualan berbasis website milik PD Indoapi Sarana Utama sehingga rekomendasi pembelian barang muncul secara otomatis dan lebih dinamis.

ACKNOWLEDGMENT

Terimakasih kepada Program Studi S1 Informatika, Fakultas Teknologi Industri dan Informatika dan LPPM IT Telkom Purwokerto untuk kesempatan mengerjakan tugas akhir dan mempublikasikannya.

DAFTAR PUSTAKA

- [1] S. Li, "A Gentle Introduction on Market Basket Analysis — Association Rules," 2017. [Online]. Available: <https://datascienceplus.com/a-gentle-introduction-on-market-basket-analysis-association-rules/>. [Accessed: 18-Dec-2017].
- [2] F. Nurchalifatun, "Penerapan Metode Asosiasi Data Mining Menggunakan Algoritma Apriori Untuk Mengetahui Kombinasi Antar Item Set Pada Pondok Kopi," in Fakultas Ilmu Komputer, Universitas Dian Nuswantoro, Semarang, 2015.
- [3] H. D. Anggraeni, R. Saputra and B. Noranita, "Aplikasi Data Mining Analisis Data Trankasi Penjualan Obat Menggunakan Algoritma Apriori," Jurnal Masyarakat Informatika, vol. IV, no. 2, pp. 1-8, 2015.
- [4] A. Ikhwan, D. Nofriansyah and Sriani, "Penerapan Data Mining Dengan Algoritma FP-Growth Untuk

- Mendukung Strategi Promosi Pendidikan," SAINTIKOM, vol. XIV(3), no. 224, pp. 1-16, 2015.
- [5] B. N. Sari, A. Rahman and Y. T. Mursityo, "Penerapan Data Mining Untuk Mengetahui Pola Asosiasi Antara Data Mahasiswa Dan Tingkat Kelulusan Menggunakan Algoritma Fold-Growth," in Universitas Brawijaya, Malang, 2015.
- [6] A. R. Mei, "Perbandingan Algoritma Apriori Dan Algoritma Fp-Growth Untuk Rekomendasi Pada Transaksi Peminjaman Buku Di Perpustakaan Universitas Dian Nuswantoro," in Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro, Semarang, Semarang, 2014.
- [7] Erwin, "Analisis Market Basket Dengan Algoritma Apriori dan FP-Growth," Jurnal Generic, vol. IV (2), no. 26, pp. 26-30, 2009. F. Nurlaela, "Sistem Pakar Untuk Mendeteksi Penyakit Gigi Pada Manusia," Indones. J. Comput. Sci., vol. 10, no. 4, pp. 76-82, 2013.