

Implementasi Algoritma Apriori dan FP-Growth Untuk Menentukan Persediaan Barang

Agus Junaidi

Fakultas Teknologi Informasi
Universitas Bina Sarana Informatika
Email: agus.asj@bsi.ac.id

Abstrak - Hasil dari penjualan pada minimarket yang selalu berupa laporan hanya dilihat tanpa ada tindak lanjut untuk menentukan keputusan diwaktu yang akan datang. Pihak manajemen hanya menerima laporan jumlah barang terjual dan berapa banyak pendapatannya. Padahal dengan data tersebut seharusnya dapat dimanfaatkan untuk menyusun rencana-rencana penjualan diwaktu yang akan datang. Dengan menggunakan metode *Frequent Pattern Growth*, pihak manajemen dapat mengambil keputusan barang mana yang membutuhkan persediaan yang lebih banyak dibandingkan dengan barang yang lain. Peletakan barang yang sesuai dengan hubungan antar barang yang biasanya dibeli konsumen juga dapat ditentukan berdasarkan *support* dan *confidence* berdasarkan market base analysis yang diperoleh dari perhitungan asosiasi dengan menggunakan metode *Frequent Pattern Growth*. Dengan menggunakan metode *Frequent Pattern Growth* maka penempatan barang dan persediaan barang pada minimarket dapat terkontrol dengan baik sehingga pelayanan pada konsumen akan meningkat yang akhirnya dapat juga meningkatkan omset penjualan. Dalam penelitian ini *support* ditentukan menggunakan ambang batas 60% dan *confidence* 90%. Dengan memperhatikan hubungan *support* dan *confidence* pemilik mini market dapat menyediakan dan menempatkan barang yang akan dijual secara tepat.

Kata Kunci — persediaan; frequent pattern growth; market base analysis

I. PENDAHULUAN

Perubahan perilaku konsumen dalam berbelanja pada toko mengharuskan pemilik toko untuk selalu melakukan inovasi dalam aspek penjualan yang dilakukannya. Permintaan konsumen yang terus meningkat secara otomatis juga harus diimbangi dengan teknologi untuk proses penjualan dan pelaporan hasil penjualannya. Selama ini, hasil dari proses penjualan hanya menjadi laporan pada perusahaan untuk mengetahui berapa barang terjual dan berapa pendapatannya. Hasil laporan penjualan hanya menjadi pemberitahuan untuk pihak manajemen dan tidak dilihat sebagai sumber data untuk pengembangan diwaktu yang akan datang, padahal hasil penjualan dapat digunakan sebagai tren untuk memprediksi permintaan barang oleh konsumen kedepannya. Dengan menggunakan metode Apriori dan Frequent Pattern Growth pihak manajemen dapat melihat *support* dan *confidence* dari barang atau produk yang telah terjual. *Support* dapat digunakan

untuk mempengaruhi penempatan barang yang biasanya dibeli oleh konsumen. Misalnya setiap konsumen yang membeli roti tawar maka juga akan membeli selai. Sedangkan *confidence* menunjukkan presentasi pembelian barang tersebut dari transaksi yang dilakukan. Dengan melihat nilai *support* dan *confidence* maka pemilik mini market dapat menempatkan barang-barang sesuai dengan permintaan konsumen. “Kekurangan atau kekosongan stok barang pada suatu toko/perusahaan akan berdampak sangat buruk untuk keberhasilan dan kelancaran transaksi jual beli, penyebab terjadinya kekosongan stok adalah tidak adanya informasi yang disampaikan dari perusahaan kepada supplier penyetok barang secara dini” [1].

Penelitian sebelumnya yang berhubungan dengan frequent pattern growth adalah “Dalam persaingan di dunia bisnis, khususnya industri Apotek, menuntut para pengembang untuk menemukan suatu strategi jitu yang dapat meningkatkan penjualan obat. Salah satu cara mengatasinya adalah dengan tetap tersediaannya berbagai jenis obat yang dibutuhkan oleh konsumen. Untuk mengetahui obat apa saja yang biasa dibeli oleh konsumen. Penerapan Algoritma Apriori dapat membantu dalam membentuk kandidat kombinasi *item*, kemudian dilakukan pengujian apakah kombinasi tersebut memenuhi parameter *support* dan *confidence* minimum yang merupakan nilai ambang yang diberikan oleh pengguna. Jika memenuhi parameter *support* dan *confidence* maka hasil tersebut dapat membantu dalam penentuan pola pembelian obat dan membantu tata letak obat berdasarkan kecenderungan konsumen membeli obat” [2].

II. LANDASAN TEORI

Dalam penelitian ini teori-teori yang digunakan adalah:

A. Data Mining

Menurut Gartner Group Data mining adalah proses menemukan korelasi baru yang bermakna, pola dan Tren oleh memilah-milah sejumlah besar data yang disimpan dalam repositori, menggunakan pola teknologi pengenalan serta teknik matematika dan statistik [3].

B. Market Basket Analysis

Market basket analysis adalah suatu metode analisa atas perilaku konsumen secara spesifik dari suatu golongan/kelompok tertentu. Sumber data dari market basket analysis antara lain dapat bersumber dari transaksi kartu kredit, kartu lotere, kupon diskon, panggilan keluhan pelanggan. Market basket analysis umumnya dimanfaatkan sebagai titik awal pencarian pengetahuan dari suatu transaksi data ketika kita tidak mengetahui pola spesifik apa yang kita cari. Kebutuhan market basket analysis berawal dari keakuratan dan manfaat yang dihasilkannya dalam wujud aturan asosiasi (association rules). Yang dimaksud dengan association rules adalah pola-pola keterkaitan data dalam basis data.[4]

Ide dasar algoritma apriori dimulai dengan mengembangkan frequent item set dengan menggunakan satu item, dan secara rekursif mengembangkan frequent item set dengan dua item, tiga item, dan seterusnya hingga frequent item set dengan semua ukuran [5].

Data mining sering mengarah pada penemuan hubungan dan korelasi diantara item pada kumpulan data transaksi atau relasional yang besar. Dengan dikumpulkan dan disimpannya data secara terus-menerus banyak industri tertarik dengan proses data mining dan database [6]

Secara umum aturan asosiasi dalam data mining mempunyai dua langkah proses yaitu: 1) mencari item-set yang sering muncul untuk menentukan minimum support, 2) menghasilkan aturan asosiasi yang kuat dari item-set tersebut, sehingga memenuhi untuk minimum support dan minimum confidence [6]

Rumus untuk menentukan Support dan Confidence adalah:

$$Support = P(A \cap B) = \frac{\text{jumlah transaksi mengandung A dan B}}{\text{total transaksi}} \dots (1)$$

$$Confidence = P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\text{jumlah transaksi mengandung A dan B}}{\text{transaksi A}} \dots (2)$$

Minimum support adalah parameter yang digunakan dalam menentukan pola dalam data mining untuk menemukan statistik dari pola-pola yang signifikan [7].

Confidence, adalah suatu ukuran yang menunjukkan hubungan antar dua item secara conditional (berdasarkan suatu kondisi tertentu) [2].

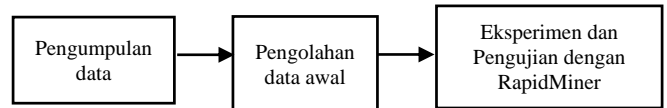
C. Algoritma FP-Growth

Algoritma FP-Growth merupakan salah satu alternatif algoritma yang cukup efektif untuk mencari himpunan data yang paling sering muncul (frequent itemset) dalam sebuah kumpulan data yang besar [8].

FP-Growth dapat menemukan frekwensi itemsets dengan hanya sedikit mengakses pada database aslinya, dan pendekatannya adalah yang paling efisien. Selain itu, FP-Growth juga dapat menghindari permasalahan jika jumlah calon itemsets-nya terlalu besar. FP-Growth menggunakan awalan khusus Tree (FP-Tree) untuk mengatur data [9].

III. METODE PENELITIAN

Penelitian dilakukan dengan menggunakan jenis eksperimen dengan tiga tahapan sebagai berikut:



Gambar 1. Tahapaan Penelitian

A. Pengumpulan data

Pengumpulan data dilakukan pada sebuah minimarket di daerah Tangerang pada bulan Juli tahun 2018. Barang dalam minimarket dikelompokkan kedalam 6 kategori yaitu Telor (Eggs), Makanan cepat saji (Fastfood), Ikan dan Hidangan laut (Fish and Seafood), Buah dan Sayur (Fruits and Vegetables), Roti (Bakery) dan Daging (Meat).

B. Pengolahan data awal

Proses pengolahan data awal dilakukan dengan mengubah data penjualan kedalam representasi biner (mengubah dengan angka 1 jika data terisi dan mengubah dengan angka 0 jika tidak terisi) dari data kategori barang yang terjual. Data hasil penjualan akan diubah kedalam format data tabular dan kemudian dicari support dan confidence.

C. Eksperimen dan Pengujian dengan RapidMiner

Setelah diubah kedalam representasi biner dalam format data tabular, maka langkah selanjutnya dilakukan eksperimen dan pengujian dengan menggunakan software RapidMiner

IV. HASIL DAN PEMBAHASAN

Hasil dari pengumpulan data penjualan pada minimarket selama satu bulan telah didapatkan data yang dapat ditampilkan dalam tabel sebagai berikut:

Tabel 1. Sampel Data

Tgl Transaksi	Eggs	Fastfood	Fish & Seafood	Fruits & Vegetables	Bakery	Meat
01-07-18	73.372	-	-	235.277	271.022	123.074
02-07-18	32.183	-	22.000	193.036	954.876	31.081
03-07-18	39.115	-	22.000	232.722	1.241.005	37.540
04-07-18	54.874	4.477	4.400	261.460	171.636	43.930
05-07-18	58.078	4.477	4.400	311.820	198.575	46.165

Tgl Transaksi	Eggs	Fastfood	Fish & Seafood	Fruits & Vegetables	Bakery	Meat
...
30-07-18	-	-	4.400	1.846.657	162.167	-
31-07-18	129.510	-	-	1.765.810	-	127.596

Dalam tabel 1 baris judul menunjukkan kelompok barang dan baris selanjutnya adalah jumlah transaksi penjualan, sedangkan kolom pertama berisi tanggal transaksi. Tahapan selanjutnya untuk dapat memasukkan data pada RapidMiner maka format data pada tabel 1 yang berupa numerik diubah menjadi format binominal yang berisi angka 1 untuk transaksi yang berisi penjualan dan angka 0 untuk transaksi yang tanpa penjualan, sehingga menjadi seperti tampilan pada tabel 2.

Tabel 2. Transaksi Penjualan dalam Format Binominal

No Urut	Eggs	Fastfood	Fish & Seafood	Fruits & Vegetables	Bakery	Meat
1	1	0	0	1	1	1
2	1	0	1	1	1	1
3	1	0	1	1	1	1
4	1	1	1	1	1	1
5	1	1	1	1	1	1
6	1	0	0	1	1	0
7	1	0	0	1	1	0
8	1	0	0	1	1	1
9	1	0	0	1	1	0
10	1	0	0	1	1	1
11	1	0	0	1	1	1
12	1	1	0	1	1	1
13	1	0	0	1	1	0
14	1	0	0	1	1	1
15	1	0	0	1	1	1
16	0	0	0	1	1	1
17	1	0	0	1	1	0
18	1	0	0	1	1	1
19	1	0	0	1	1	1
20	1	0	0	1	1	0
21	1	0	0	1	0	1
22	1	0	0	1	1	0
23	1	0	0	1	1	1
24	1	0	0	1	1	1
25	1	0	0	1	1	0
26	1	0	0	1	1	1
27	1	0	0	1	1	1

No Urut	Eggs	Fastfood	Fish & Seafood	Fruits & Vegetables	Bakery	Meat
28	1	0	1	1	1	1
29	1	0	1	0	1	1
30	0	0	1	1	1	0
31	1	0	0	1	0	1
Total	29	3	7	30	29	22

Dari format tabel 2 dapat dihitung *support* dari kategori atau gabungan kategori dengan cara menjumlahkan transaksi yang mengandung kategori tersebut dibagi dengan jumlah total transaksi tersebut, yang dapat diuraikan kedalam tabel sebagai berikut:

Tabel 3. *Support* dari setiap kategori

Kategori	Jumlah	<i>Support</i>
<i>Eggs</i>	29	94%
<i>Fastfood</i>	3	10%
<i>Fish & Seafood</i>	7	23%
<i>Fruit & Vegetables</i>	30	97%
<i>Bakery</i>	29	94%
<i>Meat</i>	22	71%

Pada tabel 3. *Support* yang mempunyai ambang batas lebih dari 60% ada empat kategori yaitu *Eggs*, *Fruit & Vegetables*, *Bakery*, dan *Meat*.

Sedangkan *support* untuk dua kategori dapat diuraikan dalam tabel 4.

Tabel 4. *Support* dari kombinasi 2 kategori

Kategori	Jumlah	<i>Support</i>
<i>Eggs, Fastfood</i>	3	10%
<i>Eggs, Fish & Seafood</i>	6	19%
<i>Eggs, Fruit & Vegetables</i>	28	90%
<i>Eggs, Bakery</i>	27	87%
<i>Eggs, Meat</i>	21	68%
<i>Fastfood, Fish & Seafood</i>	2	6%
<i>Fastfood, Fruit & Vegetables</i>	3	10%
<i>Fastfood, Bakery</i>	3	10%
<i>Fastfood, Meat</i>	3	10%
<i>Fish & Seafood, Fruit & Vegetables</i>	6	19%
<i>Fish & Seafood, Bakery</i>	7	23%
<i>Fish & Seafood, Meat</i>	6	19%
<i>Fruit & Vegetables, Bakery</i>	28	90%
<i>Fruit & Vegetables, Meat</i>	21	68%
<i>Bakery, Meat</i>	20	65%

Pada tabel 4 terlihat *support* yang melebihi ambang batas 60% ada enam kategori, yaitu: *Eggs* dan *Fruit & Vegetables*, *Eggs* dan *Bakery*, *Eggs* dan *Meat*, *Fruit & Vegetables* dan *Bakery*, *Fruit & Vegetables*, *Meat* serta *Bakery* dan *Meat*.

Tabel 5. Support dari kombinasi 3 kategori

Kategori	Jumlah	Support
Eggs, Fastfood, Fish & Seafood	2	6%
Eggs, Fastfood, Fruit & Vegetables	3	10%
Eggs, Fastfood, Bakery	3	10%
Eggs, Fastfood, Meat	3	10%
Eggs, Fish & Seafood, Fruit & Vegetables	5	16%
Eggs, Fish & Seafood, Bakery	6	19%
Eggs, Fish & Seafood, Meat	6	19%
Eggs, Fruit & Vegetables, Bakery	26	84%
Eggs, Fruit & Vegetables, Meat	20	65%
Eggs, Bakery, Meat	19	61%

Dari tabel 5, maka data kombinasi dari tiga kategori yang memenuhi ambang batas lebih dari 60% adalah tiga data yaitu: Eggs + Fruit & Vegetables + Bakery, Eggs + Fruit & Vegetables + Meat, serta Eggs + Bakery + Meat.

Langkah selanjutnya setelah menentukan support dari kombinasi kategori adalah menghitung confidence. Dengan menggunakan rumus perhitungan confidence, maka didapatkan data seperti pada tabel 6.

Tabel 6. Prosentase Confidence

No	Aturan	Confidence	
1	Eggs, Fastfood → Eggs	3/29	10%
2	Eggs, Fish & Seafood → Eggs	6/29	21%
3	Eggs, Fruit & Vegetables → Eggs	28/29	97%
4	Eggs, Bakery → Eggs	27/29	93%
5	Eggs, Meat → Eggs	21/29	72%
6	Fastfood, Fish & Seafood → Fastfood	2/3	67%
7	Fastfood, Fruit & Vegetables → Fastfood	3/3	100%
8	Fastfood, Bakery → Fastfood	3/3	100%
9	Fastfood, Meat → Fastfood	3/3	100%
10	Fish & Seafood, Fruit & Vegetables → Fish & Seafood	6/7	86%
11	Fish & Seafood, Bakery Vegetables → Fish & Seafood	7/7	100%
...
...
152	Eggs, Bakery, Meat, Fastfood, Fruits & Vegetables → Eggs, Bakery, Meat, Fastfood	3/3	100%

Pada metode FP-Growth jika mengacu pada data tabel 1, maka dapat dibuat data transaksi secara descending seperti pada tabel 7.

Tabel 7. Data transaksi per kategori untuk FP-Growth

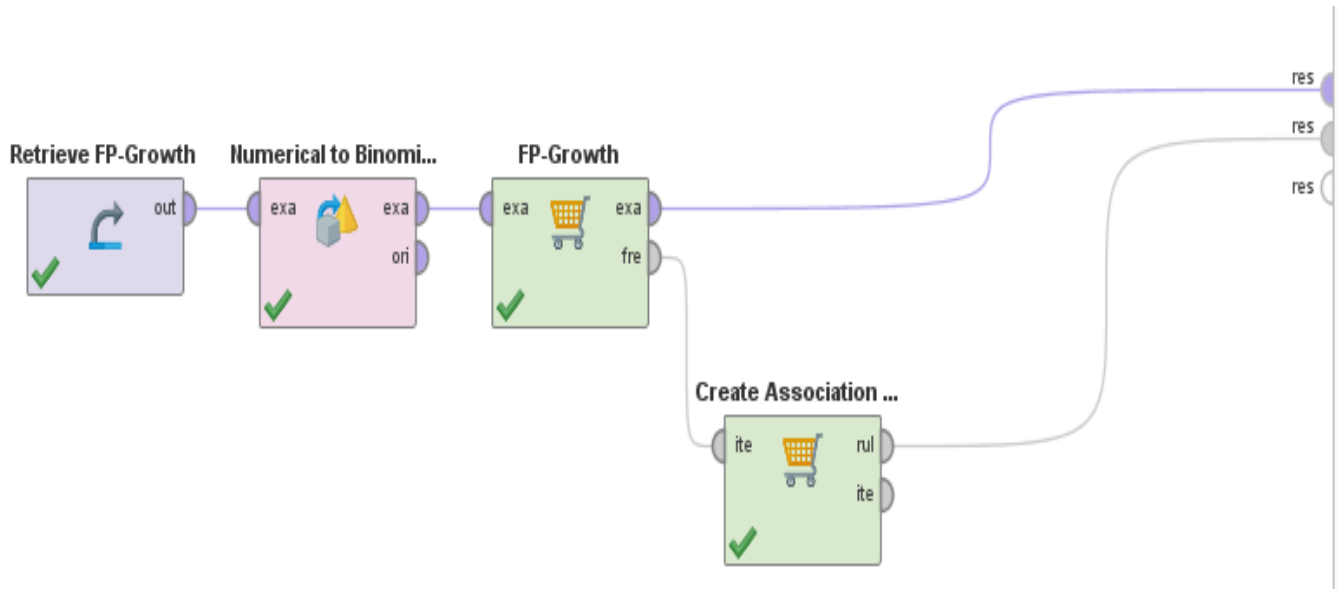
Kode	Kategori	Frekuensi
D	Fruit & Vegetables	30
A	Eggs	29
E	Bakery	29
F	Meat	22
C	Fish & Seafood	7
B	Fastfood	3

Minimum support yang ditentukan pada algoritma apriori adalah 60%, sehingga data yang memenuhi minimum support jika menggunakan metode FP-Growth adalah >=22 sehingga dapat ditampilkan hasil dari data transaksi yang memenuhi kriteria tersebut dalam tabel 8

Tabel 8. Data transaksi per kategori untuk FP-Growth yang memenuhi minimum support

No	Kategori	Minimum kategori >=22
1	{A,D,E,F}	{A,D,E,F}
2	{A,C,D,E,F}	{A,D,E,F}
3	{A,C,D,E,F}	{A,D,E,F}
4	{A,B,C,D,E,F}	{A,D,E,F}
5	{A,B,C,D,E,F}	{A,D,E,F}
6	{A,D,E}	{A,D,E}
7	{A,D,E}	{A,D,E}
8	{A,D,E,F}	{A,D,E,F}
9	{A,D,E}	{A,D,E}
10	{A,D,E,F}	{A,D,E,F}
11	{A,D,E,F}	{A,D,E,F}
12	{A,B,D,E,F}	{A,D,E,F}
13	{A,D,E}	{A,D,E}
14	{A,D,E,F}	{A,D,E,F}
15	{A,D,E,F}	{A,D,E,F}
...
...
30	{C,D,E}	{D,E}
31	{A,D,E}	{A,D,E}

Dengan menggunakan software RapidMiner 8.1 tahapan proses untuk algoritma FP-Growth dapat digambarkan sebagai berikut:



Gambar 2. FP-Growth pada RapidMiner

Row No.	Eggs	Fastfood	Fish & Seafo...	Fruits & Veg...	Bakery	Meat
20	true	false	false	true	true	false
21	true	false	false	true	false	true
22	true	false	false	true	true	false
23	true	false	false	true	true	true
24	true	false	false	true	true	true
25	true	false	false	true	true	false
26	true	false	false	true	true	true
27	true	false	false	true	true	true
28	true	false	true	true	true	true
29	true	false	true	false	true	true
30	false	false	true	true	true	false
31	true	false	false	true	false	true

Gambar 3. Proses Numerical to Binominal dari sampel data penjualan


```

Association Rules
[Meat] --> [Fruits & Vegetables, Eggs, Bakery] (confidence: 0.818)
[Fruits & Vegetables, Fish & Seafood] --> [Eggs] (confidence: 0.833)
[Eggs, Fish & Seafood] --> [Fruits & Vegetables] (confidence: 0.833)
[Fruits & Vegetables, Fish & Seafood] --> [Meat] (confidence: 0.833)
[Meat, Fish & Seafood] --> [Fruits & Vegetables] (confidence: 0.833)
[Fruits & Vegetables, Fish & Seafood] --> [Eggs, Bakery] (confidence: 0.833)
[Eggs, Fish & Seafood] --> [Fruits & Vegetables, Bakery] (confidence: 0.833)
[Fruits & Vegetables, Bakery, Fish & Seafood] --> [Eggs] (confidence: 0.833)
[Eggs, Bakery, Fish & Seafood] --> [Fruits & Vegetables] (confidence: 0.833)
[Fruits & Vegetables, Fish & Seafood] --> [Eggs, Meat] (confidence: 0.833)
[Eggs, Fish & Seafood] --> [Fruits & Vegetables, Meat] (confidence: 0.833)
[Meat, Fish & Seafood] --> [Fruits & Vegetables, Eggs] (confidence: 0.833)
[Eggs, Meat, Fish & Seafood] --> [Fruits & Vegetables] (confidence: 0.833)
[Fruits & Vegetables, Fish & Seafood] --> [Bakery, Meat] (confidence: 0.833)
[Fruits & Vegetables, Bakery, Fish & Seafood] --> [Meat] (confidence: 0.833)
[Meat, Fish & Seafood] --> [Fruits & Vegetables, Bakery] (confidence: 0.833)
[Bakery, Meat, Fish & Seafood] --> [Fruits & Vegetables] (confidence: 0.833)
[Fruits & Vegetables, Fish & Seafood] --> [Eggs, Bakery, Meat] (confidence: 0.833)
[Eggs, Fish & Seafood] --> [Fruits & Vegetables, Bakery, Meat] (confidence: 0.833)
[Fruits & Vegetables, Bakery, Fish & Seafood] --> [Eggs, Meat] (confidence: 0.833)
[Eggs, Bakery, Fish & Seafood] --> [Fruits & Vegetables, Meat] (confidence: 0.833)
[Meat, Fish & Seafood] --> [Fruits & Vegetables, Eggs, Bakery] (confidence: 0.833)
[Eggs, Meat, Fish & Seafood] --> [Fruits & Vegetables, Bakery] (confidence: 0.833)
[Bakery, Meat, Fish & Seafood] --> [Fruits & Vegetables, Eggs] (confidence: 0.833)

```

Gambar 6. Association Rules dari RapidMiner

Dari pengolahan data penjualan pada gambar 6 yang dilakukan proses perhitungan menggunakan RapidMiner didapatkan jumlah aturan asosiasi sebanyak 152 *rule* dengan support terendah adalah 10% dan tertinggi 94%, sedangkan *confidence* terendah adalah 10% dan tertinggi adalah 100%.

V. KESIMPULAN

Berdasarkan data dan hasil pembahasan maka dapat diambil kesimpulan sebagai berikut:

- Secara keseluruhan dari data sampel penjualan diperoleh 152 *rule* yang terdiri dari 24 *rule* asosiasi yang memenuhi support dengan ambang batas 60% dan 108 *rule* yang memenuhi *confidence* 90%.
- Penggunaan algoritma apriori dengan metode FP-Growth dapat membantu pihak manajemen untuk meletakkan barang yang biasa dibeli oleh konsumen, sehingga memudahkan pelanggan untuk dapat membeli barang tersebut.
- Metode FP-Growth dapat membantu pihak manajemen untuk memantau stok barang yang sering dibeli oleh konsumen sehingga tidak akan terjadi kelangkaan pasokan.

DAFTAR PUSTAKA

- [1] M. Sholik and A. Salam, "Implementasi Algoritma Apriori untuk Mencari Asosiasi Barang yang Dijual di E-commerce OrderMas," *Techno.com*, vol. 17, no. 2, pp. 158–170, 2018.
- [2] R. K. Robi Yanto, "Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat," *Implementasi Data Min. dengan Metod. Algoritm. Apriori dalam Menentukan Pola Pembelian Obat, Citec J.*, vol. Vol. 2, No, no. ISSN: 2354-5771, pp. 102–113, 2015.
- [3] D. T. Larose and C. D. Larose, *Discovering Knowledge in Data*. 2014.
- [4] G. Gunadi and D. I. Sensuse, "Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (Fp-Growth).," *Telematika*, vol. 4, no. 1, pp. 118–132, 2012.
- [5] B. Santosa and A. Umam, *Data Mining dan Big Data Analytics*, 2nd ed. Yogyakarta: Penebar Media Pustaka, 2018.
- [6] J. Han, M. Kamber, and J. Pei, "Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems)," 2011.
- [7] A. Salam and M. S. H. Khayal, "Mining top-k frequent patterns without minimum support threshold," *Knowl. Inf. Syst.*, vol. 30, no. 1, pp. 57–86, 2012.
- [8] F. Fatihatul, A. Setiawan, and R. Rosadi, "Asosiasi Data Mining Menggunakan Algoritma FP-Growth Untuk Market Basket Analysis," *Jatinangor Univ. Padjadjaran*, pp. 1–8, 2011.
- [9] B. Wang *et al.*, "Comprehensive Association Rules Mining of Health Examination Data with an Extended FP-Growth Method," *Mob. Networks Appl.*, vol. 22, no. 2, pp. 267–274, 2017.