

PERBANDINGAN AKURASI ALGORITMA APRIORI DAN FP-GROWTH UNTUK MENGETAHUI POLA PENYEBAB GANGGUAN INDIHOME AREA WONOSOBO

¹⁾Irgi Sena Fajar Lukerti, ²⁾Muhamad Fuat Asnawi, ³⁾Nulngafan

^{1,2)}Universitas Sains Al-Qur'an

¹⁾irgi2403@gmail.com, ²⁾fuatasnawi@unsiq.ac.id, ³⁾affan@unsiq.ac.id

INFO ARTIKEL

Riwayat Artikel :

Diterima : 17 November 2024

Disetujui : 29 November 2024

Kata Kunci :

Association Rules, Data Mining, Apriori, FP-Growth, Indihome.

ABSTRAK

Seiring berjalananya waktu, pelanggan Indihome terus bertambah banyak dan tidak menutup kemungkinan terjadi gangguan pada layanan Indihome. Untuk mengatasi masalah tersebut, perlu dilakukan analisis terhadap data yang ada, sehingga dapat diketahui pola-pola penyebab gangguan pada layanan Indihome area Wonosobo. Metode Association Rules digunakan untuk menemukan korelasi menarik dari kumpulan data yang besar, kemudian dilakukan perbandingan hasil asosiasi menggunakan algoritma Apriori dan FP-Growth. Data gangguan Indihome yang digunakan berjumlah 458 data, dengan menggunakan minimum support 5% dan minimum confidence 50%, algoritma Apriori mendapatkan persentase 17,65% rule asosiasi yang valid sedangkan algoritma FP-Growth mendapatkan persentase 100% rule asosiasi yang valid. Hal ini menunjukkan bahwa algoritma FP-Growth lebih direkomendasikan untuk menganalisa pola penyebab gangguan Indihome area Wonosobo, karena pengetahuan yang diperoleh lebih akurat.

ARTICLE INFO

Article History :

Received : Nov 17, 2024

Accepted : Nov 29, 2024

Keywords:

Association Rules, Data Mining, Apriori, FP-Growth, Indihome.

ABSTRACT

As time goes by, Indihome customers continue to grow and it does not rule out the possibility of interference with Indihome services. To overcome this problem, it is necessary to analyze the existing data, so that patterns of causes of interference in Indihome services in the Wonosobo area can be identified. The Association Rules method is used to find interesting correlations from large data sets, then compare the association results using the Apriori and FP-Growth algorithms. The Indihome interference data used amounted to 458 data, using a minimum support of 5% and a minimum confidence of 50%, the Apriori algorithm got a percentage of 17.65% valid association rules while the FP-Growth algorithm got a percentage of 100% valid association rules. This shows that the FP-Growth algorithm is more recommended for analyzing patterns of causes of Indihome interference in the Wonosobo area, because the knowledge obtained is more accurate.

1. PENDAHULUAN

PT Telkom Indonesia (Persero) Tbk (Telkom) adalah Badan Usaha Milik Negara (BUMN) yang bergerak di bidang jasa layanan teknologi informasi dan komunikasi (TIK) dan jaringan telekomunikasi di Indonesia. Telkom mulai saat ini membagi bisnisnya menjadi 3 Digital Business Domain, diantaranya yaitu Digital Connectivity, Digital Platform dan Digital Services. Salah satu produk Telkom dari Digital Connectivity adalah Indihome (Indonesia Digital Home) merupakan layanan telekomunikasi data yang terdiri dari Telepon Rumah, Internet, dan IPTV menggunakan jaringan Fiber Optic.

Sampai dengan Desember 2022, terhitung pelanggan Indihome area Wonosobo mencapai 2485 pelanggan (sumber: kpro.telkom.co.id/kpro/pstelkom) dan pada bulan Januari sampai Desember 2022 jumlah gangguan Indihome area Wonosobo mencapai 458 pelanggan (sumber: datum.tr4.my.id). Hal ini tentunya sangat mengganggu aktivitas sehari-hari dan dapat menyebabkan kerugian bagi pelanggan dan juga penyedia layanan. Untuk mengatasi masalah tersebut, perlu dilakukan analisis terhadap data yang ada, sehingga dapat diketahui pola-pola penyebab gangguan pada layanan Indihome area Wonosobo. Salah satu metode analisis data yang dapat digunakan adalah dengan menggunakan algoritma data mining, khususnya algoritma Apriori dan FP-Growth.

Data Mining adalah suatu pencarian dan analisa dari jumlah data yang sangat besar dan bertujuan untuk mencari arti dari pola dan aturan (Chien, Wang, & Cheng, 2007). Pengertian lain menyebutkan bahwa Data Mining adalah proses menemukan pola yang menarik dan pengetahuan dari data yang berjumlah besar (Septiani, 2017). Algoritma Apriori dan FP-Growth merupakan bagian dari metode association rule, yaitu teknik data mining untuk menemukan aturan suatu kombinasi antar item.

Algoritma Apriori adalah algoritma untuk menemukan pola frekuensi tinggi (Aldino, 2021). Mempunyai kelebihan lebih sederhana, dapat menangani data yang besar (dalam penggunaan memori saat memproses data yang besar tentunya berpengaruh terhadap banyaknya item yang diproses) serta mudah dipahami

struktur kerja dan implementasinya (Fauzy, 2015).

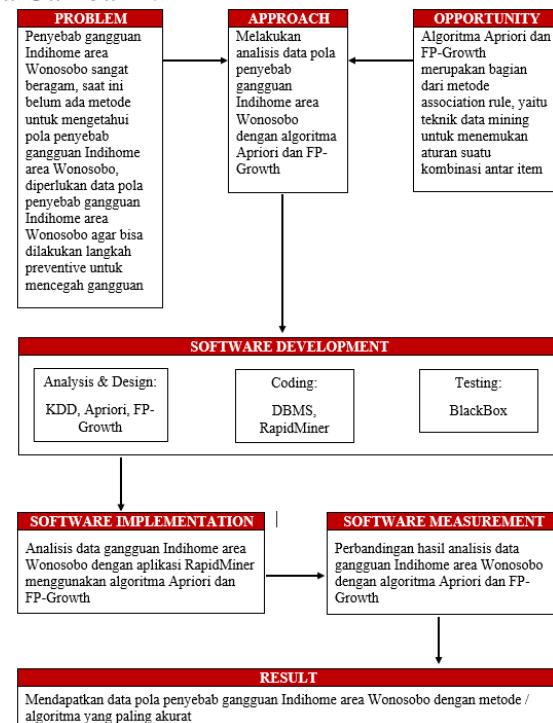
Algoritma FP-Growth (Frequent Pattern-Growth) adalah salah satu algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (Anas, 2022). Walaupun algoritma FP-Growth turunan dari algoritma apriori, namun pendekatannya berbeda dengan paradigma yang digunakan pada metode apriori. FP-Growth membangun FP-Tree untuk mengompres database transaksi yang asli (Elvira, 2021).

Penelitian ini bertujuan untuk mengetahui pola penyebab gangguan Indihome area Wonosobo dengan Algoritma Apriori dan Algoritma FP-Growth serta mengetahui perbandingan akurasi Algoritma Apriori dan FP-Growth dalam analisis pola penyebab gangguan Indihome area Wonosobo menggunakan aplikasi RapidMiner. Setelah mengetahui pola penyebab gangguan Indihome, diharapkan bisa menjadi langkah preventif yang tepat sasaran untuk mencegah terjadinya gangguan Indihome di area Wonosobo.

2. METODE

2.1. Kerangka Pikir

Kerangka pikir penelitian ini ditunjukkan pada Gambar 1.



Gambar 1. Kerangka Pikir

2.2. Metode Pengumpulan Data

Pengumpulan data dilakukan di kantor PT Telkom Akses Wonosobo dengan metode pengumpulan data yaitu observasi, dan studi pustaka. Observasi dilakukan dengan pengumpulan data informasi melalui pengamatan langsung pada objek penelitian, sehingga diperoleh masalah-masalah yang ada. Studi Pustaka dilakukan dengan mencari dan mempelajari jurnal penelitian terdahulu yang berkaitan dengan data mining terutama penggunaan metode algoritma Apriori dan FP-Growth. Dari hasil pengumpulan data tersebut diperoleh data jumlah gangguan Indihome pada bulan Januari - Desember 2022 sebanyak 458 gangguan.

2.3. Metode Pengolahan Data

Metode pengolahan data pada penelitian ini dilakukan berdasarkan proses KDD (Knowledge Discovery in Database). Berikut penjelasan masing-masing proses di dalam KDD (Knowledge Discovery in Database):

a. Selection

Data yang sudah didapat dilakukan proses pemilihan terlebih dahulu. Data yang digunakan adalah data gangguan indihome area Wonosobo selama 1 tahun dari bulan Januari 2022 sampai Desember 2022.

b. Preprocessing / Cleaning

Sebelum proses data mining, perlu dilakukan proses cleaning pada data yang menjadi fokus KDD. Proses cleaning mencakup antara lain membuang duplikasi data, memeriksa data yang tidak konsisten dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi).

c. Transformation

Transformation adalah mengubah data ke dalam model analitis serta memodelkan data agar sesuai dengan analisa yang diharapkan dan format data yang diperlukan oleh algoritma data mining. Data harus diubah sebelum proses data mining. Data diubah menjadi tabel tabular, sehingga data tersebut dapat diolah selanjutnya dengan menggunakan software RapidMiner.

d. Data Mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Proses pengolahan data didasarkan

pada algoritma yang sesuai dengan teknik data mining menggunakan software RapidMiner. Algoritma yang digunakan pada penelitian ini adalah algoritma Apriori dan FP-Growth

e. Interpretation / Evaluation

Tahap ini merupakan hasil dari proses data mining berupa pola-pola yang akan diuji pada hipotesa yang sudah dibuat sebelumnya. Sehingga akan memperoleh kesimpulan-kesimpulan yang mendekati hasil atau hipotesa untuk proses selanjutnya.

3. HASIL DAN PEMBAHASAN

Setelah melalui proses KDD (Knowledge Discovery in Database) diperoleh data kombinasi itemset, frekuensi, support dan confidence pada Tabel 1 dibawah. Dimana F = Frekuensi, S = Support dan C = Confidence.

Tabel 1 Kombinasi Itemset

Kombinasi	F	S	C
[offline, kabel luar putus]	197	43,01%	57,94%
[online, konfigurasi hilang]	74	16,16%	62,71%
[offline, modem rusak]	48	10,48%	14,12%
[offline, odp kotor / rusak]	29	6,33%	8,53%
[offline, kabel dalam putus]	26	5,68%	7,65%
[online, stb rusak]	26	5,68%	22,03%
[offline, gangguan massal]	19	4,15%	5,59%
[offline, konektor kotor]	14	3,06%	4,12%
[online, edukasi pelanggan]	10	2,18%	8,47%
[online, remote stb rusak]	5	1,09%	4,24%
[offline, kabel terjepit]	4	0,87%	1,18%
[offline, adaptor modem rusak]	3	0,66%	0,88%
[online, port lan mati]	3	0,66%	2,54%

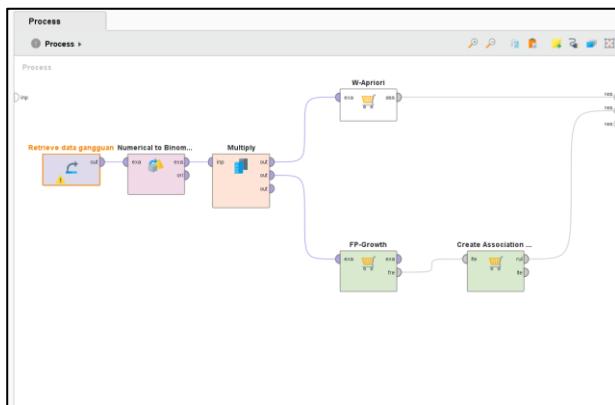
Untuk menentukan nilai support sebuah item dapat diperoleh dengan menggunakan rumus persamaan (1) seperti di bawah ini :

$$\text{Support } A \cap B = \frac{\text{Jumlah transaksi mengandung } A \text{ dan } B}{\text{Total Transaksi}} \quad (1)$$

Sedangkan untuk menentukan nilai confidence dapat ditentukan dengan rumus persamaan (2) seperti dibawah ini :

$$\text{Confidence } A \rightarrow B = \frac{\text{Jumlah transaksi mengandung } A \text{ dan } B}{\text{Jumlah transaksi mengandung } A} \quad (2)$$

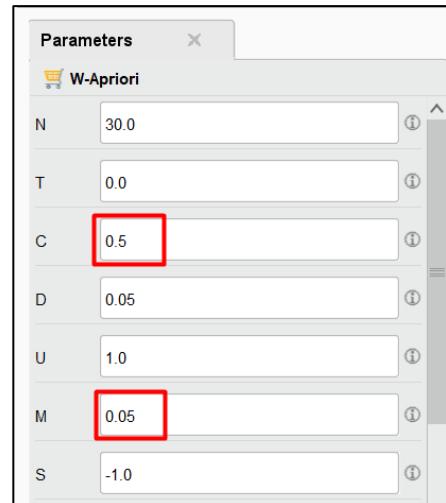
Berdasarkan perhitungan diatas, maka penulis menentukan nilai minimum support sebesar 5% dan nilai minimum confidence sebesar 50%.



Gambar 2 Desain Proses Algoritma Apriori dan FP-Growth dengan RapidMiner

3.1. Hasil Data Mining Algoritma Apriori

Pada Gambar 3 dibawah terlihat jendela parameter W-Apriori, yang digunakan untuk menentukan minimum support dan minimum confidence. Kolom C adalah minimum confidence, diisi dengan nilai 0.5 (50%) sedangkan kolom M adalah minimum support diisi dengan nilai 0.05 (50%).



Gambar 3 Parameter Algoritma Apriori

```

W-Apriori
=====
Apriori
=====

Minimum support: 0.05 (23 instances)
Minimum metric <confidence>: 0.5
Number of cycles performed: 19

Generated sets of large itemsets:
Size of set of large itemsets L(1): 8
Size of set of large itemsets L(2): 5
Size of set of large itemsets L(3): 2

Best rules found:
1. online=true 118 ==> offline=false 118 conf:(1)
2. offline=false 118 ==> online=true 118 conf:(1)
3. konfigurasi hilang=true 74 ==> offline=false 74 conf:(1)
4. konfigurasi hilang=true 74 ==> online=true 74 conf:(1)
5. online=true konfigurasi hilang=true 74 ==> offline=false 74 conf:(1)
6. offline=false konfigurasi hilang=true 74 ==> online=true 74 conf:(1)
7. konfigurasi hilang=true 74 ==> offline=false online=true 74 conf:(1)
8. stb rusak=true 26 ==> offline=false 26 conf:(1)
9. stb rusak=true 26 ==> online=true 26 conf:(1)
10. online=true stb rusak=true 26 ==> offline=false 26 conf:(1)
11. offline=false stb rusak=true 26 ==> online=true 26 conf:(1)
12. stb rusak=true 26 ==> offline=false online=true 26 conf:(1)
13. offline=false 118 ==> konfigurasi hilang=true 74 conf:(0.63)
14. online=true 118 ==> konfigurasi hilang=true 74 conf:(0.63)
15. offline=false online=true 118 ==> konfigurasi hilang=true 74 conf:(0.63)
16. online=true 118 ==> offline=false konfigurasi hilang=true 74 conf:(0.63)
17. offline=false 118 ==> online=true konfigurasi hilang=true 74 conf:(0.63)

```

Gambar 4 Aturan Asosiasi yang dihasilkan Algoritma Apriori

Pada Gambar 4 merupakan hasil dari proses Data Mining dengan algoritma Apriori menggunakan aplikasi RapidMiner, terdapat 17 aturan asosiasi yang dihasilkan, yaitu :

- Rule 1: $\text{online} = \text{true } 118 \rightarrow \text{offline} = \text{false } 118 \text{ conf:}(1)$
- Rule 2: $\text{offline} = \text{false } 118 \rightarrow \text{online} = \text{true } 118 \text{ conf:}(1)$
- Rule 3: $\text{konfigurasi hilang} = \text{true } 74 \rightarrow \text{offline} = \text{false } 74 \text{ conf:}(1)$
- Rule 4: $\text{konfigurasi hilang} = \text{true } 74 \rightarrow \text{online} = \text{true } 74 \text{ conf:}(1)$
- Rule 5: $\text{online} = \text{true } \text{konfigurasi hilang} = \text{true } 74 \rightarrow \text{offline} = \text{false } 74 \text{ conf:}(1)$

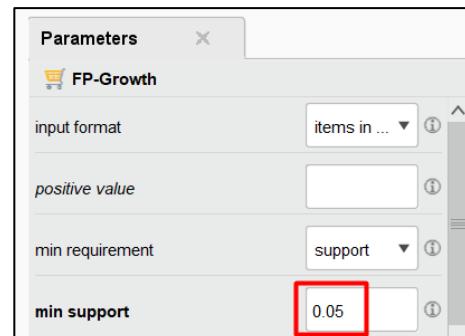
- f. Rule 6: offline = false konfigurasi hilang = true 74 → online = true 74 conf:(1)
- g. Rule 7: konfigurasi hilang = true 74 → offline = false online = true 74 conf:(1)
- h. Rule 8: stb rusak = true 26 → offline = false 26 conf:(1)
- i. Rule 9: stb rusak = true 26 → online = true 26 conf:(1)
- j. Rule 10: online = true stb rusak = true 26 → offline = false 26 conf:(1)
- k. Rule 11: offline = false stb rusak = true 26 → online = true 26 conf:(1)
- l. Rule 12: stb rusak = true 26 → offline = false online = true 26 conf:(1)
- m. Rule 13: offline = false 118 → konfigurasi hilang = true 74 conf:(0.63)
- n. Rule 14: online = true 118 → konfigurasi hilang = true 74 conf:(0.63)
- o. Rule 15: offline = false online = true 118 → konfigurasi hilang = true 74 conf:(0.63)
- p. Rule 16: online = true 118 → offline = false konfigurasi hilang = true 74 conf:(0.63)
- q. Rule 17: offline = false 118 → online = true konfigurasi hilang = true 74 conf:(0.63)

Dari 17 aturan asosiasi yang didapat, ada 14 aturan asosiasi yang tidak valid (tidak bisa ditarik kesimpulan) karena ada variabel bernilai false, jadi hanya ada 3 aturan asosiasi yang valid (bisa ditarik kesimpulan) yaitu :

- a. Rule 4: Jika penyebab gangguan adalah konfigurasi hilang maka jenis gangguan adalah online, nilai confidence 100%.
- b. Rule 9: Jika penyebab gangguan adalah stb rusak maka jenis gangguan adalah online, dengan nilai confidence 100%.
- c. Rule 14: Jika jenis gangguan adalah online, maka 63% kemungkinan penyebab gangguan adalah konfigurasi hilang.

3.2. Hasil Data Mining Algoritma FP-Growth

Sama seperti algoritma Apriori dalam proses data mining di aplikasi RapidMiner, langkah pertama adalah menentukan nilai minimum support sebesar 0.05 (5%) pada operator FP-Growth dan menentukan nilai minimum confidence sebesar 0.5 (50%) pada operator Create Association Rules. Terlihat pada Gambar 5 dan Gambar 6.



Gambar 5 Menentukan nilai minimum support



Gambar 6 Menentukan nilai minimum confidence

No.	Premises	Conclusion	Support	Confidence
1	offline	kabel luar putus	0.430	0.579
2	online	konfigurasi hilang	0.162	0.627

Gambar 7 Aturan Asosiasi yang dihasilkan Algoritma Apriori dalam bentuk tabel

Association Rules				
[offline] --> [kabel luar putus] (confidence: 0.579)				
[online] --> [konfigurasi hilang] (confidence: 0.627)				

Gambar 8 Aturan Asosiasi yang dihasilkan Algoritma Apriori dalam bentuk deskripsi

- Pada Gambar 7 dan Gambar 8 merupakan hasil dari proses Data Mining dengan algoritma Apriori menggunakan aplikasi RapidMiner, terdapat 2 aturan asosiasi yang dihasilkan, yaitu:
- a. Rule 1: Jika jenis gangguan adalah offline, maka 57% kemungkinan penyebab gangguan adalah kabel luar putus, dengan didukung oleh 43% dari data keseluruhan.
 - b. Rule 2: Jika jenis gangguan adalah online, maka 63% kemungkinan penyebab gangguan adalah konfigurasi hilang, dengan didukung oleh 16% dari data keseluruhan.

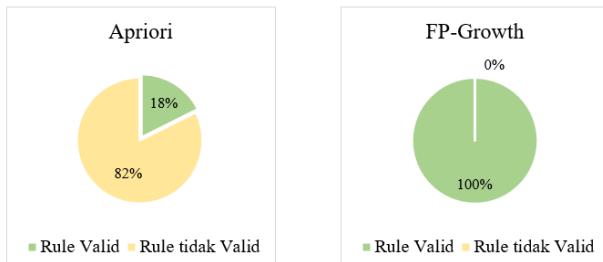
3.3. Hasil Perbandingan Algoritma Apriori dan FP-Growth

Setelah dilakukan pengujian pada algoritma Apriori dan FP-Growth, hasil perbandingan dari

kedua algoritma tersebut dapat dilihat pada Tabel 2 dibawah:

Tabel 2 Hasil perbandingan Algoritma Apriori dan FP-Growth

Algoritma	Jumlah Itemset	Aturan Asosiasi	Akurasi
Apriori	2 kombinasi (contoh: online → konfigurasi hilang) dan 3 kombinasi (contoh: online, offline → konfigurasi hilang)	17	18%
FP-Growth	2 kombinasi saja (contoh: offline → kabel luar putus)	2	100%



Gambar 9 Grafik perbandingan Algoritma Apriori dan FP-Growth

Berdasarkan hasil perbandingan dari Tabel 2 dan Gambar 9 diatas, pengujian dengan algoritma Apriori menghasilkan 2 kombinasi dan 3 kombinasi yang seharusnya hanya 2 kombinasi saja, sementara algoritma FP-Growth menghasilkan 2 kombinasi saja, yang artinya sudah sesuai dengan ketentuan. Kemudian untuk akurasi pada algoritma Apriori bernilai 18% dan algoritma FP-Growth bernilai 100%. Maka dalam konteks data gangguan Indihome area Wonosobo, algoritma FP-Growth lebih direkomendasikan untuk menganalisa pola penyebab gangguan Indihome area Wonosobo, karena pengetahuan yang diperoleh lebih akurat.

4. PENUTUP

4.1. Kesimpulan

Penelitian terhadap 458 data gangguan Indihome di area Wonosobo selama Januari hingga Desember 2022 mengungkap pola

penyebab gangguan, yaitu: gangguan offline 57% kemungkinan disebabkan oleh kabel luar putus (43% support), sedangkan gangguan online 63% kemungkinan disebabkan oleh konfigurasi hilang (16% support). Analisis perbandingan algoritma Apriori dan FP-Growth menunjukkan bahwa algoritma FP-Growth lebih direkomendasikan untuk menganalisis pola penyebab gangguan tersebut dengan parameter minimum support dan confidence yang sama.

4.2. Saran

Penelitian ini masih memiliki kekurangan, sehingga disarankan agar penelitian selanjutnya menggunakan data selain gangguan, seperti data pemasangan baru atau data material, serta memanfaatkan aplikasi lain seperti Weka atau Tanagra untuk membandingkan hasil analisis dari berbagai aplikasi.

5. DAFTAR PUSTAKA

- Aldino, A. A., Pratiwi, E. D., Sintaro, S., & Putra, A. D. (2021, October). Comparison of market basket analysis to determine consumer purchasing patterns using fp-growth and apriori algorithm. In 2021 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE) (pp. 29-34). IEEE.
- Anas, S., Rumui, N., Roy, A., & Saputro, P. H. (2022). Comparison of apriori algorithm and fp-growth in managing store transaction data. International Journal of Computer and Information System (IJCIS), 3(4), 158-162.
- Destiyati, O. S. A., & Aribowo, E. (2015). Analisis Perbandingan Algoritma Apriori Dan Algoritma Hash Based Pada Market Basket Analysis Di Apotek UAD. Jurnal Sarjana Teknik Informatika, 3(1), 1–10.
- Istiqomah, A. D., Astuti, Y., & Nurjanah, S. (2022). Implementasi Algoritma FP-Growth dan Apriori Untuk Persediaan Produk. Jurnal Informatika Polinema, 8(2), 37–42.
- lucyana hutahaean, elfrida, Safii, M., & Efendi Damanik, B. (2020). Implementasi algoritma apriori pada sistem persediaan barang. JIKO (Jurnal Informatika Dan Komputer), 3(3), 173–180. <https://doi.org/10.33387/jiko.v3i3.2192>

- Maryani, I., Revianti, O., Nur, H. M., & Sunanto, S. (2022). Implementasi Data Mining Pada Penjualan Di Toko GOC Kosmetik Dengan Menggunakan Metode Algoritma Apriori. *Indonesian Journal on Software Engineering (IJSE)*, 8(1), 92–98. <https://doi.org/10.31294/ijse.v8i1.13017>
- Maulidiya, H., & Jananto, A. (2020). Asosiasi Data Mining Menggunakan Algoritma Apriori dan FP-Growth sebagai Dasar Pertimbangan Penentuan Paket Sembako. *Proceeding SENDIU 2020*, 6, 36–42.
- Musdalifah, I., & Jananto, A. (2022). Analisis Perbandingan Algoritma Apriori Dan FP-Growth Dalam Pembentukan Pola Asosiasi Keranjang Belanja Pelanggan. *Progresif: Jurnal Ilmiah Komputer*, 18(2), 175. <https://doi.org/10.35889/progresif.v18i2.878>
- Nurasiah. (2021). Implementasi Algoritma FP-Growth Pada Pengenalan Pola Penjualan. *TIN : Terapan Informatika Nusantara*, 1(9), 438–444.
- Oktavia Gama, A. W., Gede Darma Putra, I. K., & Agung Bayupati, I. P. (2016). Implementasi Algoritma Apriori Untuk Menemukan Frequent Itemset Dalam Keranjang Belanja. *Majalah Ilmiah Teknologi Elektro*, 15(2), 21–26. <https://doi.org/10.24843/mite.1502.04>
- Pahlevi, O., Sugandi, A., & Sintawati, I. D. (2018). Penerapan Algoritma Apriori dalam Pengendalian Kualitas Produk. *Jurnal & Penelitian Teknik Informatika*, 3(1), 272–278.
- Sianturi, F. A. (2018). Penerapan Algoritma Apriori Untuk Penentuan Tingkat Pesanan. *Jurnal Mantik Penusa*, 2(1), 50–57. <http://bowmasbow.blogspot.com/20>
- Yudho Ardianto, M., & Adinugroho, S. (2021). Penentuan Tata Letak Produk menggunakan Algoritma FP-Growth pada Toko ATK. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 5(9), 3826–3832. <http://j-ptiik.ub.ac.id>.