

## AN ANALYSIS OF THE CORRELATION OF CONSUMER PURCHASE PATTERNS USING APRIORI METHOD (A CASE STUDY OF CV. VIFOR CIPTA SOLUSI)

**Junianto Bagas Prasetyo<sup>\*</sup>, Mutaqin Akbar**

Program Studi Informatika, Fakultas Teknologi Informasi

Universitas Mercu Buana Yogyakarta

Jl. Wates Km. 10 Yogyakarta 55753, Indonesia

Email Corresponding: <sup>\*</sup>zerogi26@gmail.com

Received: November 2, 2020; Accepted : November 8, 2020; Published : November 14, 2020

### ***Abstract***

*Cv. Vifor Cipta Solusi Is A Company That provides electronic equipment. The competition in business, especially among companies providing electronic equipment, is growing. To increase sales of the products, business people in this field need to set a strategy. This research applied the data mining with association technique to find out strong combinations between items. Apriori algorithm is part of the association method in data mining which aims to find out frequent item sets from a certain set of data. The apriori algorithm process is carried out by determining frequent itemsets that meet the requirements of the predetermined minimum support and minimum confidence. From these two measurements, the final association rule can be obtained by which the value of the accuracy (validity) of the association is calculated using the lift ratio. This research obtained 2 association rules with item attributes totaling 294 types of products from 20 main attributes with 1813 records processed from 5282 raw transaction data from January to December 2019. The strong rules obtained are MagnetToner HP Laserjet 85 A, Drum Toner HP Laserjet 85 A => HP Laserjet 85 A toner with a support value of 3.42%, and a confidence value of 81.58% with a lift ratio of 5.00. If a consumer buys an HP Laserjet 85 A MagnetToner and a HP Laserjet 85 A Toner Drum, the consumer will also buy an HP Laserjet 85 A Toner.*

**Keyword:** data mining, apriori algorithm, electronic stores

### **1. Latar Belakang Masalah**

CV. Vifor Cipta Solusi adalah perusahaan yang bergerak di bidang penyedia alat elektronik. Persaingan dalam bisnis khususnya perusahaan yang bergerak di bidang penyedia alat elektronik semakin banyak. Agar dapat meningkatkan penjualan produk yang dijual, para pelaku bisnis di bidang ini harus mempunyai strategi yang tepat guna mengambil langkah ke depannya. Salah satu cara yang bisa dilakukan oleh perusahaan dengan cara memanfaatkan data dari transaksi penjualan yang ada.

Dengan menggunakan data transaksi tersebut, data dapat diolah lebih lanjut sehingga didapatkan informasi baru dari kumpulan data yang ada. Produk pada perusahaan yang ditawarkan bermacam dengan berbagai macam produk serta berbagai tipe, dan terkadang merk juga mempengaruhi masyarakat untuk membeli produk tersebut. Untuk mengetahui produk dengan penjualan terbanyak dan keterkaitan produk satu dengan produk yang lainnya

tetap diperlukan, serta pentingnya tetap menjaga tersedianya stock berbagai barang yang dibutuhkan oleh konsumen.

*Data Mining* adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basisdata dengan melakukan penggalian pola pola dari data dengan tujuan untuk memanipulasi data menjadi informasi yang lebih berharga yang diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat dalam basisdata [1]. Untuk mengetahui kebiasaan pelanggan atau perilaku pelanggan mengenai barang apa saja yang sering dibeli. Dengan penerapan Algoritma Apriori dapat membantu dalam membentuk kandidat kombinasi dari *Item produk*, kemudian dilakukan pengujian apakah kombinasi tersebut memenuhi parameter *support* dan *confidence minimum* yang merupakan batas ambang nilai yang diberikan oleh pengguna. Jika memenuhi parameter *support* dan *confidence* maka hasil tersebut dapat sebagai acuan dalam membantu pihak perusahaan dalam manajemen perencanaan, seperti penentuan pola *Item-set* pembelian barang serta membantu tata letak barang berdasarkan kecenderungan konsumen membeli barang. Algoritma apriori dapat digunakan untuk mengetahui kebiasaan atau perilaku pelanggan seperti penerapannya pada data penjualan [2][3][4][5][6], data persediaan [7][8].

## 2. Metodologi Penelitian

Untuk mengetahui produk penjualan terbanyak yang terjual secara bersamaan dibutuhkan sebuah metode agar didapatkan informasi baru. Yaitu dengan menggunakan algoritma Apriori dan bantuan *tools* Tanagra. Algoritma Apriori adalah jenis aturan asosiasi pada *Data Mining* [9]. Dengan menambang data atau upaya untuk menggali informasi dan pengetahuan yang berharga pada *database* yang sangat besar disebut *Data Mining* atau *Knowledge Discovery in Database* disingkat KDD. Salah satu algoritma yang paling populer pada teknik *Data Mining* adalah algoritma Apriori. Sedangkan dalam penemuan pola kombinasi hubungan antar *Item-sets* digunakan *Association rules* (Aturan Asosiasi) [10]. Dalam jurnalnya kemudahan untuk menjual suatu produk di pasaran menggunakan media internet membuat banyaknya pedagang pedagang baru bermunculan. Keterbatasan modal menjadi kendala sebagian besar pedagang baru tersebut. Penerapan algoritma apriori untuk mencari produk yang laku di pasaran dapat mengatasi masalah pedagang baru tersebut. Algoritma apriori merupakan salah satu cabang dari *Data Mining*, yang berfungsi untuk mencari informasi informasi dari sebuah kumpulan data. Dalam penelitian ini, hasil algoritma apriori dapat digunakan untuk menjadi bahan pertimbangan dalam membuat strategi pemasaran, pengoptimalan penjualan, serta sebagai bahan pertimbangan untuk melakukan penyetokan ulang.

### 2.1 Prinsip Kerja Sistem

Analisis asosiasi atau *Association rule* mining adalah teknik *Data Mining* untuk menemukan aturan asosiatif antara suatu kombinasi *Item*. Metodologi dasar analisis asosiasi terbagi menjadi dua tahap :

#### 1. Scanning

Tahap ini mencari kombinasi *Item* yang memenuhi syarat *minimum* dari nilai *support* dalam *database*. Nilai *support* sebuah *Item* diperoleh dengan rumus sebagai berikut [11]:

$$\text{Support (A)} = \frac{\text{Jumlah Transaksi mengandung A}}{\text{Total transaksi}} \times 100\% \quad (1)$$

Nilai *support* dari 2 *Item* diperoleh dengan menggunakan rumus berikut:

$$\text{Support (A, B)} = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi}} \times 100\% \quad (2)$$

*Frequent Itemset* menunjukkan *Itemset* yang memiliki frekuensi kemunculan lebih dari nilai *minimum* yang ditentukan. Misalkan nilai *minimum* = 2, maka semua *itemsets* yang frekuensi kemunculannya lebih dari atau sama dengan 2 kali disebut *frequent*.

## 2. Pembentukan Aturan Asosiasi

Setelah semua pola pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat *minimum* untuk *confidence* dengan menghitung *confidence* aturan asosiatif  $A \rightarrow B$ . Nilai *Confidence* dari aturan  $A \rightarrow B$  diperoleh rumus berikut.

$$Confidence = P(B|A) = \frac{\sum \text{Transaksi mengandung } A \text{ dan } B}{\sum \text{Transaksi mengandung } A} \quad (3)$$

Untuk menentukan aturan asosiasi yang akan dipilih maka harus diurutkan berdasarkan  $Support \times Confidence$ . Aturan diambil sebanyak n-aturan yang memiliki hasil terbesar.

Nilai korelasi dapat diketahui dengan menggunakan rumus persamaan sebagai berikut:

$$Lift(A, B) = \frac{P(A \cup B)}{P(A)P(B)} \quad (4)$$

Apabila dari perhitungan tersebut menghasilkan nilai dibawah 1 maka terdapat korelasi negatif. Untuk perhitungan yang menghasilkan nilai diatas 1 maka terdapat korelasi positif. Namun apabila menghasilkan nilai sama dengan 1 maka tidak ada korelasi antara X dan Y.

## 3. Hasil dan Pembahasan

Data satu tahun terakhir lebih tepatnya tanggal 01 Januari 2019 – 31 Desember 2019 sejumlah 5286 data mentah. Setelah melakukan proses data *pre-processing* (data *cleaning*) dilakukan perbaikan ditemukan kesalahan penulisan, *duplikasi* data atau *inkonsistensi* pada data transaksi sehingga data yang di dapatkan adalah 1813 data transaksi *clean* setelah di seleksi dan yang dipakai adalah 294 jenis produk dari 20 atribut, dengan format *upload* data bertipe .CSV agar dapat dibaca oleh sistem yang sudah dirancang. Penelitian ini menghasilkan hasil analisis berupa tabel yang berisi kombinasi dari nilai *support* dan *confidence*. Nilai *support* dari semua atribut *Item* berbeda dapat kita peroleh dari rumus persamaan 1. Daftar hasil perhitungan nilai *Support* dari tiap atribut seperti terlihat pada Tabel 1:

Tabel 1 Lolos *Support Minimum*.

Sup	Item	Jml	Sup	Ket
2%	Cartridge Epson PLQ 20	212	11,69	Lolos
	Printer HP Laserjet 1132	12	0,66	Tidak Lolos
3,2%	Kabel Cat 6	296	16,33	Lolos
	Toner HP Laserjet 85 A	143	7,89	Lolos

Hasil kombinasi 2-*itemset* yang lolos dapat dilihat pada Tabel 2.

Tabel 2 Hasil kombinasi 2-*itemset* yang lolos

Min Sup	Item		Jml	Sup	Ket
2%	Cartridge Epson PLQ 20	MagnetToner HP Laserjet 85 A	12	0,66	Tidak Lolos
	Cartridge Epson PLQ 20	Drum Toner HP Laserjet 85 A	12	0,66	Tidak Lolos
3,2%	MagnetToner HP Laserjet 85 A	Kamera CCTV Hikvision IP Cam	4	0,22	Tidak Lolos

Min Sup	Item		Jml	Sup	Ket
	MagnetToner HP Laserjet 85 A	Kamera CCTV Outdoor Hikvision	3	0,17	Tidak Lolos

Hasil kombinasi 3-Itemset yang terbentuk dapat kita peroleh sebagai perhitungan persamaan rumus 2 dapat dilihat pada Tabel 3.

Tabel 3 Hasil kombinasi 3-itemset yang lolos

Min Sup	Item			Jml	Sup	Ket
2%	Cartridge Epson PLQ 20	Toner HP Laserjet 85 A	Cartridge Olivetty PR 2E	4	0,22	Tidak Lolos
	Cartridge Epson PLQ 20	Toner HP Laserjet 85 A	Doctor BladeToner HP Laserjet 85 A	2	0,11	Tidak Lolos
3,2%	Toner HP Laserjet 85 A	MagnetToner HP Laserjet 85 A	Drum Toner HP Laserjet 85 A	62	3,42	Lolos
	Toner HP Laserjet 85 A	Toner HP Laserjet 83 A	MagnetToner HP Laserjet 85 A	45	2,48	Tidak Lolos

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat *minimum* untuk *Confidence* dengan ketentuan 80% mengacu pada persamaan rumus 3. Selanjutnya menghitung nilai *minimum* nilai *confidence* dari 2-itemset dan 3-itemset yang terbentuk. Hasil dari *confidence* dari calon aturan asosiasi seperti tampak pada Tabel 4.

Tabel 4 Hasil perhitungan *minimum confidence*

Sup	Con	Itemset	X => Y	Sup X U Y	Sup X	Min con	Ket
2%	70%	itemset 2	Drum Toner HP Laserjet 85 A => MagnetToner HP Laserjet 85 A	4,19	9,32	44,97	Tidak Lolos
		itemset 3	Toner HP Laserjet 83 A => Drum Toner HP Laserjet 83 A, MagnetToner HP Laserjet 83 A	2,10	7,94	26,39	Tidak Lolos
			MagnetToner HP Laserjet 83 A, Drum Toner HP Laserjet 83 A => Toner HP Laserjet 83 A	2,10	2,32	90,48	Lolos
2%	80%	itemset 2	Drum Toner HP Laserjet 83 A => MagnetToner HP Laserjet 83 A	2,32	4,36	53,16	Tidak Lolos
			MagnetToner HP Laserjet 83 A => Drum Toner HP Laserjet 83 A	2,32	4,03	57,53	Tidak Lolos
		itemset 3	MagnetToner HP Laserjet 83 A => Toner HP Laserjet 83 A, Drum Toner HP Laserjet 83 A	2,10	4,03	52,05	Tidak Lolos
			Toner HP Laserjet 83 A => Drum Toner HP Laserjet 83 A, MagnetToner HP Laserjet 83 A	2,10	7,94	26,39	Tidak Lolos
3,2%	70%	itemset 2	Drum Toner HP Laserjet 83 A => Toner HP Laserjet 83 A	3,42	4,36	78,48	Lolos

Sup	Con	Itemset	X => Y	Sup X U Y	Sup X	Min con	Ket
			MagnetToner HP Laserjet 85 A => Drum Toner HP Laserjet 85 A	4,19	7,89	53,15	Tidak Lolos
		itemset 3	Toner HP Laserjet 85 A, MagnetToner HP Laserjet 85 A => Drum Toner HP Laserjet 85 A	3,42	6,40	53,45	Tidak Lolos
			MagnetToner HP Laserjet 85 A, Drum Toner HP Laserjet 85 A => Toner HP Laserjet 85 A	3,42	4,19	81,58	Lolos
3,2%	81%	itemset 2	Drum Toner HP Laserjet 85 A => Toner HP Laserjet 85 A	7,12	9,32	76,33	Tidak Lolos
			MagnetToner HP Laserjet 85 A => Toner HP Laserjet 85 A	6,40	7,89	81,12	Lolos
		itemset 3	MagnetToner HP Laserjet 85 A => Toner HP Laserjet 85 A, Drum Toner HP Laserjet 85 A	3,42	7,89	43,36	Tidak Lolos
			Drum Toner HP Laserjet 85 A => MagnetToner HP Laserjet 85 A, Toner HP Laserjet 85 A	3,42	9,32	36,69	Tidak Lolos

Tabel 5 Nilai *Lift Ratio 2-itemset*

No	X => Y	Confidence	Nilai Uji lift	Korelasi <i>rule</i>
1	MagnetToner HP Laserjet 85 A => Toner HP Laserjet 85 A	81,12	$\frac{81,12}{16,32}$ = 4,97	korelasi positif

Tabel 6 Nilai *Lift Ratio 3-itemset*

No	X => Y	Confidence	Nilai Uji lift	Korelasi <i>rule</i>
1	MagnetToner HP Laserjet 85 A, Drum Toner HP Laserjet 85 A => Toner HP Laserjet 85 A	81,58	$\frac{81,58}{16,32}$ = 5,00	korelasi positif

Dari hasil penelitian data pada Tabel 5 dan Tabel 6, dapat disimpulkan bahwa pola pembelian di Cv.Vifor Cipta Solution dengan aturan asosiasi, dengan nilai *support* sebesar 3,2%, *confidence* sebesar 81%, dan *lift ratio* sebesar *lift*>1. Analisis dilakukan pada 1813 *record* transaksi menghasilkan 2 *rule* asosiasi yang terbentuk. Dengan hasil analisa *rule* yang terbentuk seperti Tabel 7.

Tabel 7 Hasil analisa

No	Hasil
1	Jika konsumen membeli MagnetToner HP Laserjet 85 A, Drum Toner HP Laserjet 85 A, maka konsumen juga akan membeli Toner HP Laserjet 85 A Dengan Persentase 81,58%
2	Jika konsumen membeli MagnetToner HP Laserjet 85 A, maka konsumen juga akan membeli Toner HP Laserjet 85 A Dengan Persentase 81,12%

#### 4. Kesimpulan

Pola asosiasi yang terbentuk dengan nilai *minimum support* 3,2% dan nilai *minimum confidence* 81% menghasilkan 2 aturan asosiasi. Dan *strong rules* yang didapatkan adalah MagnetToner HP Laserjet 85 A, Drum Toner HP Laserjet 85 A => Toner HP Laserjet 85 A dengan nilai *support* 3,42% dan nilai *confidence* 81,58%.

Penerapan Algoritma Apriori dalam penelitian ini menghasilkan pola pembelian. Jika konsumen membeli MagnetToner HP Laserjet 85 A, Drum Toner HP Laserjet 85 A, maka konsumen juga akan membeli Toner HP Laserjet 85 A dan Jika konsumen membeli MagnetToner HP Laserjet 85 A, maka konsumen juga akan membeli Toner HP Laserjet 85 A.

Analisis pola yang dihasilkan dari aplikasi ini dapat dijadikan sebagai acuan pengambilan keputusan pihak Cv.Vifor Cipta Solusi dalam membuat promosi serta rekomendasi kombinasi tata letak *item*. Rekomendasi produk dari hasil aplikasi ini yang bisa dijadikan sebagai promosi bagi perusahaan yaitu: Paket 3 *item* MagnetToner HP Laserjet 85 A, Drum Toner HP Laserjet 85 A dan Toner HP Laserjet 85 A dan Paket 2 *item* MagnetToner HP Laserjet 85 A dengan Toner HP Laserjet 85 A.

Penelitian selanjutnya dapat dikembangkan menggunakan metode yang berbeda, salah satunya yaitu *fp-growth* dan dapat dibandingkan nilai parameter *support* dan *confidence* mana yang lebih baik.

#### Daftar Pustaka

- [1] Ristianingrum, R., & Sulastri, S. (n.d.). Implementasi Data Mining Menggunakan Algoritma Apriori. Sintak 2017, 372–382.
- [2] Aditya, A., Marisa, F., & Purnomo, D. (2016). Penerapan Algoritma Apriori Terhadap Data Penjualan di Toko Gudang BM. Journal of Information Technology and Computer Science, 1(1), 1–5.
- [3] Fajar Rodiansyah, S. (2015). Algoritma Apriori untuk Analisis Keranjang Belanja pada Data Transaksi Penjualan. Infotech Journal, 1(2), 36–39.
- [4] Rahmawati, F., & Merlina, N. (2018). Metode Data Mining Terhadap Data Penjualan Sparepart Mesin Fotocopy Menggunakan Algoritma Apriori. Jurnal Penelitian Ilmu Komputer, System Embedded & Logic, 6(1), 9–20.
- [5] Tana, M. P., Marisa, F., & Wijaya, I. D. (2018). Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Pada Toko Oase Menggunakan Algoritma Apriori. Jurnal Informatika Merdeka Pasuruan, 3(2), 17–22.
- [6] Yanto, R., & Khoiriah, R. (2015). Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat. Citec Journal, 2(2), 102–113.
- [7] Simbolon, P. H. (2019). Implementasi Data Mining Pada Sistem Persediaan Barang Menggunakan Algoritma Apriori (Studi Kasus: Srikandi Cash Credit Elektronik dan Furniture). Jurnal Riset Komputer, 6(4), 401–406.
- [8] Tampubolon, K., Saragih, H., & Reza, B. (2013). Implementasi data mining algoritma apriori pada sistem persediaan alat-alat kesehatan. Informasi Dan Teknologi Ilmiah, 1(1).
- [9] Han, J., Pei, J., Kamber, M., & Safari, an O. M. C. (2011). Data Mining: Concepts and Techniques, 3rd Edition.
- [10] Santoso, L. W. (2003). Pembuatan Perangkat Lunak Data Mining Untuk Penggalian Kaidah Asosiasi Menggunakan Metode. Jurnal Informatika, 4(2), 49–56.
- [11] Listriani, D., Setyaningrum, A. H., & Eka, F. (2018). Penerapan Metode Assosiasi Menggunakan Algoritma Pada Aplikasi Analisa Pola Belanja Konsumen (Studi Kasus Toko Buku Gramedia Bintaro). Jurnal Teknik Informatika, 9(2). <https://doi.org/10.15408/jti.v9i2.5602>