

Penerapan Data Mining Asosiasi pada Pola Transaksi dengan Metode Apriori

Nita Syahputri

Universitas Potensi Utama, Medan, Sumatera Utara
nieta20d@gmail.com

Abstract

Indonesian business is a promising thing, the competition that occurs in the business of business people is always trying to be strategic and breakthroughs that can ensure the continuity of what they do. The youngest business cafe / restaurant, especially among young people to adults. The development of the Café / Restaurant business in Indonesia is growing rapidly, this can be seen from the many cafes / restaurants that have emerged which provide attractive interior design venues and offer a comfortable and pleasant atmosphere to gather with family and friends. Many competitors in business, especially in cafes / restaurants, require developers to find strategies that can increase sales, one of which is the use of transaction data. Data Mining is the process of extracting information from data sets through the use of algorithms and techniques involving the fields of statistics, machine learning, and database management system. A priori algorithms are a type of association rule in data mining. In this study using secondary data. The results of calculations and analysis of consumer research using Data Mining with the Apriori Algorithm method, overall from the sales sample data obtained 52 association rules that meet support above 20% and 50% confidence. Based on the test results using Rapidminer, it can be rejected by that the product that consumers are more interested in is the minimum confidence above 50%.

Keywords: Datamining, Apriori, Cafe, Sales, Pematangsiantar

Abstrak

Di Indonesia bisnis merupakan hal yang paling menjanjikan, persaingan yang terjadi dalam dunia bisnis memaksa para pelaku bisnis untuk selalu memikirkan strategis-strategis dan terobosan yang dapat menjamin kelangsungan dari bisnis yang dijalankannya. Terutama bisnis Café/Restoran paling diminati kalangan anak muda hingga dewasa. Perkembangan bisnis Café/Restoran di Indonesia berkembang sangat pesat, hal ini dapat dilihat dari banyaknya Café/Restoran yang muncul dengan menyediakan tempat berdesain interior yang menarik serta menawarkan suasana nyaman dan menyenangkan untuk berkumpul bersama Keluarga dan sahabat. Banyaknya pesaing dalam bisnis khususnya pada Café/Restoran menuntut para pengembang untuk menemukannya strategi yang dapat meningkatkan penjualan, salah satunya dengan pemanfaatan data transaksi Data Mining adalah proses ekstraksi informasi dari kumpulan data melalui penggunaan algoritma dan teknik yang melibatkan bidang ilmu statistik, mesin pembelajaran, dan sistem manajemen database. Algoritma apriori termasuk jenis aturan asosiasi pada data mining. Pada penelitian ini menggunakan data sekunder. Hasil perhitungan dan analisis penelitian konsumen menggunakan Data Mining dengan metode Algoritma Apriori, secara keseluruhan dari data sampel penjualan diperoleh 52 rule asosiasi yang memenuhi support diatas 20% dan confidence 50%. Berdasarkan hasil pengujian dengan menggunakan Rapidminer dapat disimpulkan bahwa produk yang lebih banyak diminati oleh konsumen adalah minimum confidence diatas 50%.

Kata Kunci: Datamining, Apriori, Cafe, Penjualan, Pematangsiantar

1. PENDAHULUAN

Di Indonesia bisnis merupakan hal yang paling menjanjikan, persaingan yang terjadi dalam dunia bisnis memaksa para pelaku bisnis untuk selalu memikirkan strategis-strategis dan terobosan yang dapat menjamin kelangsungan dari bisnis yang dijalkannya. Terutama bisnis Café/Restoran paling diminati kalangan anak muda hingga dewasa. Perkembangan bisnis Café/Restoran di Indonesia berkembang sangat pesat, hal ini dapat dilihat dari banyaknya Café/Restoran yang muncul dengan menyediakan tempat berdesain interior yang menarik serta menawarkan suasana nyaman dan menyenangkan untuk berkumpul bersama Keluarga dan sahabat. Banyaknya pesaing dalam bisnis khususnya pada Café/Restoran menuntut para pengembang untuk menemukannya strategi yang dapat meningkatkan penjualan, salah satunya dengan pemanfaatan data transaksi [1]. Dengan jumlah transaksi yang cukup besar, perusahaan membutuhkan alat bantu analisis untuk memberikan informasi yang bermanfaat bagi perusahaan dalam penentuan [2] seperti makanan atau minuman apa yang paling diminati oleh konsumen. Maka dari ini perlu di bangun sebuah sistem untuk menentukan dan mengetahui banyaknya makanan atau minuman yang diminati oleh konsumen. Dimana hal ini sangat berguna bagi konsumen dalam menentukan makanan dan minuman yang enak.

Banyak cabang ilmu komputer yang dapat menyelesaikan masalah secara kompleks antara lain: datamining [3][4]. Datamining adalah proses yang menggunakan teknik statistic, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [5]. Data mining juga dikenal dengan istilah pattern recognition merupakan suatu algoritma yang digunakan untuk pengolahan data guna menemukan pola yang tersembunyi dari data yang diolah. Data yang diolah kemudian menghasilkan suatu pengetahuan baru yang bersumber dari data lama, hasil dari pengolahan data tersebut dapat digunakan dalam menentukan keputusan di masa depan [3]. Peneliti menggunakan cabang ilmu datamining dengan metode Apriori untuk menyelesaikan masalah transaksi penjualan DL Café & Resto. Ada beberapa penelitian terdahulu yang berkaitan dengan apriori dalam menyelesaikan permasalahan. Salah satunya penelitian yang dilakukan [2]. Hasil dari penelitian tersebut menyatakan bahwa metode apriori dapat diterapkan, secara keseluruhan dari data sampel penjualan diperoleh 142 rule asosiasi yang memenuhi support diatas 10% dan confidence 50%. Berdasarkan hasil pengujian dengan menggunakan Rapidminer dapat disimpulkan bahwa produk yang lebih banyak diminati oleh konsumen adalah minimum confidence diatas 50%. Selain itu alasan penulis menggunakan apriori adalah pengimplemantasian yang mudah di jalankan, proses yang dibutuhkan relatif singkat, sangat fleksibel dan mudah digunakan, serta menggunakan prinsip sederhana yang dijelaskan dalam non-statistik. Diharapkan hasil penelitian ini dapat membantu para

konsumen serta owner dalam menentukan makanan dan minuman yang paling diminati.

2. METODOLOGI PENELITIAN

2.1. Data Mining

Data mining adalah proses yang menggunakan teknik statistic, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [6]. Proses dari data mining yaitu mencari pola data yang tertentu menggunakan metode untuk mengolahnya. Istilah data mining dan Knowledge Discovery in Database (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar [7]. Data mining sendiri memiliki kemampuan untuk mengolah data dari suatu basis data agar menghasilkan sebuah informasi baru yang bermanfaat.

2.2. Metode Apriori

Association Rule atau Algoritma apriori merupakan salah satu algoritma yang pada dasarnya untuk menemukan Frequentitemsets (kelompok setiap item-item yang memenuhi syarat minimum support) dari setiap data transaksi [8]. Apriori digunakan untuk menentukan pola *association rule* yang tepat dan akurat serta dapat memproses data dalam jumlah yang besar dan menyeleksinya menjadi beberapa *rule*.

Langkah-langkah dalam algoritma apriori [9] :

- a) Analisis pola frekuensi tinggi (*frequent item set*), yaitu mencari kombinasi item yang memenuhi syarat minimum dari nilai *support* dalam data transaksi.

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

Adapun nilai *support* dari dua item diperoleh dengan rumus sebagai berikut.

$$Support(A \cap B) = \frac{\text{jumlah transaksi mengandung } A \text{ dan } B}{\text{Total transaksi}} \times 100\% \quad (2)$$

Secara umum, nilai *support* juga dapat dicari dengan rumus :

$$Support = \frac{\text{jumlah item yang dibeli}}{\text{jumlah seluruh transaksi}} \times 100\% \quad (3)$$

- b) Pembentukan aturan asosiatif, untuk memenuhi syarat minimum *confidence*.

$$Confidence(A, B) = \frac{\text{jumlah transaksi mengandung } A \text{ dan } B}{\text{jumlah transaksi mengandung } A} \times 100\% \quad (4)$$

Secara umum, nilai *confidence* juga dapat dicari menggunakan rumus :

$$Confidence = \frac{\text{jumlah item yang dibeli}}{\text{jumlah transaksi pada bagian antecedent}} \times 100\% \quad (5)$$

2.3. Sumber Data

Pada penelitian ini, sumber data yang diperoleh dengan melakukan observasi dan pengumpulan data yang bersumber dari DL Café and Resto yang berlokasi di Jalan Kyai Ahmad Dahlan, Kecamatan Sitalasari Bukit Sofa Pematangsiantar.

3. HASIL DAN PEMBAHASAN

3.1. Rancangan Penelitian

Proses analisis data merupakan teknik yang dilakukan secara wawancara dan proses dokumentasi atau catatan transkrip penjualan mulai dari tanggal 21 sampai 27 Januari 2021. Setelah hasil yang didapatkan menggunakan bantuan *software Microsoft Excel* akan diaplikasikan ke *RapidMiner* untuk menyesuaikan hasil yang didapatkan. Adapun tahapan yang dilakukan pada rancangan penelitian ini sebagai berikut :

a) Mengambil data dari dataset excel, seperti pada tabel 3.1 berikut :

Tabel 1. Tabel Transaksi Data Penjualan

Tanggal	Produk													
	MG	NG	NAP	BS	NIN	LT	TM	KH	CPC	J	FF	IC	SB	SM
21 januari 2021	1	1	0	0	0	1	0	0	0	1	1	0	1	0
22 januari 2021	0	0	1	0	1	0	1	1	0	0	1	1	0	0
23 januari 2021	0	0	0	1	0	0	0	0	1	1	1	0	1	0
24 januari 2021	1	0	1	0	1	1	1	0	0	1	0	1	0	1
25 januari 2021	0	1	0	1	0	1	0	0	0	0	1	0	0	0
26 januari 2021	0	1	0	1	1	1	1	0	1	1	1	0	0	1
27 januari 2021	1	1	0	1	0	1	1	0	0	1	1	1	1	0

b) Pembentuk 1 item set

Proses Pembentukan C1 atau disebut dengan 1 itemset dengan jumlah minimum support 10 dan Confidence 50%

$$Support (A) = \frac{jumlah\ transaksi\ mengandung\ A}{Total\ transaksi} \times 100\%$$

c) Analisis Data

Pada penelitian yang dilakukan penulis, data yang digunakan adalah data penjualan DL Café & Resto pada tanggal 21-27 januari 2021. Kemudian data tersebut dijadikan kedalam beberapa set item.

$$Support (Mie Goreng) = \frac{jumlah\ transaksi\ Mie\ Goreng}{Total\ transaksi} \times 100\% \\ = \frac{3}{7} \times 100\% = 42,86\%$$

Tabel 2. Support itemset-1

No	Menu	Banyak Transaksi	Support %
1	Mie Goreng	3	42,86%
2	Nasi Goreng	4	57,14%
3	Nasi Ayam Penyet	2	28,57%
4	Beef steak	4	57,14%

No	Menu	Banyak Transaksi	Support %
5	Nasi Ikan Nila	3	42,86%
6	Lemon Tea	5	71,43%
7	Teh Manis	4	57,14%
8	Kopi Hitam	1	14,29%
9	Cappucino	2	28,57%
10	Jus	5	71,43%
11	French Fries	6	85,71%
12	Ice Cream	3	42,86%
13	Salad Buah	3	42,86%
14	Smoothies	2	28,57%

$$\begin{aligned}
 \text{Support } (Mg, Ng) &= \frac{\text{jumlah transaksi } Mg, Ng}{\text{Total transaksi}} \times 100\% \\
 &= \frac{2}{7} \times 100\% = 28,57\%
 \end{aligned}$$

Tabel 3. Support itemset-2

No	Pasangan Item	Banyak Transaksi	Support %
1	Mg,Ng	2	28.57%
2	Mg,NAP	1	14.29%
3	Mg,Bs	1	14.29%
4	Mg,Nin	1	14.29%
...
83	J,Ic	2	28.57%
84	J,Sb	3	42.86%
85	J,Sm	1	14.29%
86	Ff,Ic	2	28.57%
87	Ff,Sb	3	42.86%
88	Ff,Sm	1	14.29%
89	Ic, Sb	1	14.29%
90	Ic,Sm	1	14.29%
91	Sb,Sm	0	0.00%

Nilai support tiga item dalam penelitian ini dapat dilihat pada tabel dibawah ini :

$$\begin{aligned}
 \text{Support} &= \frac{\text{jumlah item yang dibeli}}{\text{jumlah seluruh transaksi}} \times 100\% \\
 &= \frac{0}{3} \times 100\% = 0,00\%
 \end{aligned}$$

Tabel 4. Support itemset-3

No	Pasangan Item	Banyak Transaksi	Support %
1	Mg,Ng,Nap	0	0.00%
2	Mg,Ng,Bs	0	0.00%
3	Mg,Ng,Nin	0	0.00%

No	Pasangan Item	Banyak Transaksi	Support %
4	Mg,Ng,Lt	2	28.57%
5	Mg,Ng,Tm	1	14.29%
6	Mg,Ng,Kh	0	0.00%
7	Mg,Ng,Cpc	0	0.00%
...
353	cpc,ff,sm	1	14.29%
354	cpc,ic,sb	0	0.00%
355	cpc,ic,sm	0	0.00%
356	cpc,sb,sm	0	0.00%
357	j,ff,ic	1	14.29%
358	j,ff,sb	3	42.86%
359	j,ff,sm	1	14.29%
360	j,ic,sb	1	14.29%
361	j,ic,sm	1	14.29%
362	j,sb,sm	0	0.00%
363	ff,ic,sb	1	14.29%
364	ff,ic,sm	0	0.00%
365	ff,sb,sm	0	0.00%
366	ic,sb,sm	0	0.00%

Karena kombinasi 3 itemset tidak ada yang memenuhi minimal Support, maka 2 kombinasi yang memenuhi untuk pembentukan asosiasi.

$$\begin{aligned}
 \text{Confidence} &= \frac{\text{jumlah item yang dibeli}}{\text{jumlah transaksi pada bagian antecedent}} \times 100\% \\
 &= \frac{2}{3} \times 100\% = 0,00\%
 \end{aligned}$$

Tabel 5. Hasil Perhitungan *Confidence*

No	Pasangan Item	Banyak Transaksi	Confidence %
1	Mg,Ng	2	66.67%
2	Mg,NAP	1	33.33%
3	Mg,Bs	1	33.33%
4	Mg,Nin	1	33.33%
5	Mg,Lt	3	100.00%
6	Mg,Tm	2	66.67%
7	Mg,Kh	0	0.00%
8	Mg,Cpc	0	0.00%
9	Mg,J	3	100.00%
10	Mg,FF	2	66.67%
...
79	Cpc,Ic	0	0.00%
80	Cpc,Sb	1	50.00%
81	Cpc,Sm	1	50.00%

No	Pasangan Item	Banyak Transaksi	Confidence %
82	J,Ff	4	80.00%
83	J,Ic	2	40.00%
84	J,Sb	3	60.00%
85	J,Sm	1	20.00%
86	Ff,Ic	2	33.33%
87	Ff,Sb	3	50.00%
88	Ff,Sm	1	16.67%
89	Ic, Sb	1	33.33%
90	Ic,Sm	1	33.33%
91	Sb,Sm	0	0.00%

Minimum confidence yang ditentukan adalah 50%, maka item-item yang memiliki nilai confidence kurang dari 50% akan dihilangkan seperti tabel dibawah ini.

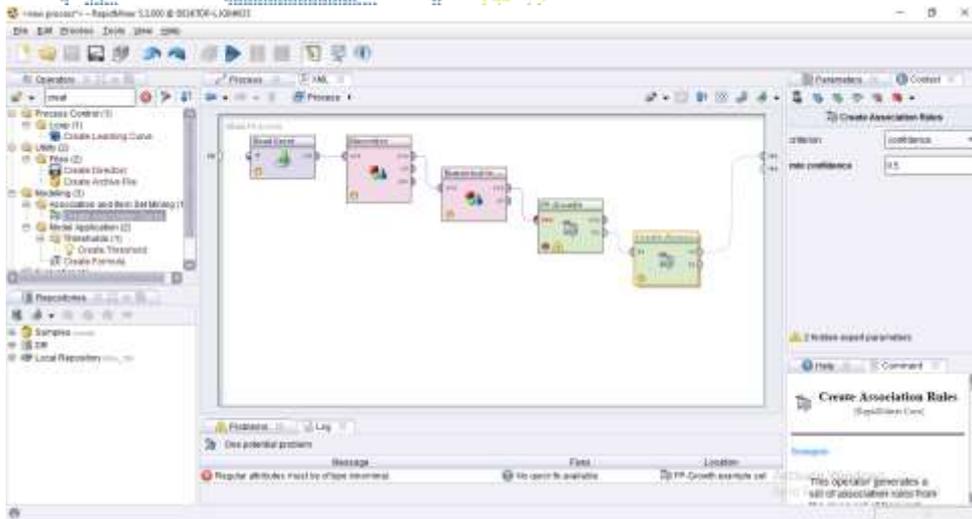
Tabel 6. Minimum Confidence

No	Pasangan Item	Banyak Transaksi	Confidence %
1	Mg,Ng	2	66.67%
2	Mg,Lt	3	100.00%
3	Mg,Tm	2	66.67%
4	Mg,J	3	100.00%
5	Mg,FF	2	66.67%
6	Mg,Ic	2	66.67%
...
45	Cpc,Ff	2	100.00%
46	Cpc,Sb	1	50.00%
47	Cpc,Sm	1	50.00%
48	J,Ff	4	80.00%
49	Cpc,J	2	100.00%
50	Cpc,Ff	2	100.00%
51	J,Sb	3	60.00%
52	Ff,Sb	3	50.00%

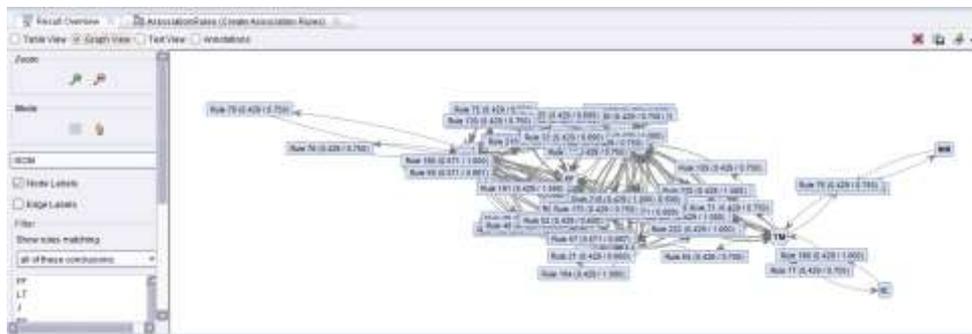
Hasilnya adalah terbentuk 52 aturan asosiasi hasil proses berdasarkan parameter yang telah ditentukan yaitu minimum support 20% dan minimum confidence 50%. Pada salah satu aturan yang terbentuk, misalnya aturan : "Mie goreng (Mg) => Lemon Tea (Lt)" dengan nilai confidence 100% berarti bahwa 100% dari konsumen yang membeli "Mie goreng juga membeli Lemon Tea".

3.2. Implementasi

Berisi hasil implementasi ataupun pengujian.



Gambar 1. Proses Pemasukan Data Melalui Software Rapid Miner



Gambar 2. Hasil Graph View Melalui Software Rapid Miner

The screenshot shows the 'Table View' of the association rules. The table lists 19 rules with their IDs, confidence values, and support values. The 'Confidence' column is sorted in descending order.

Id	Confidence	Support
11	0.429	0.008
12	0.429	0.008
1	0.429	0.008
17	0.429	0.008
18	0.429	0.008
19	0.429	0.008
13	0.429	0.008
14	0.429	0.008
15	0.429	0.008
16	0.429	0.008
2	0.429	0.008
3	0.429	0.008
4	0.429	0.008
5	0.429	0.008
6	0.429	0.008
7	0.429	0.008
8	0.429	0.008
9	0.429	0.008
10	0.429	0.008
16	0.429	0.008
17	0.429	0.008
18	0.429	0.008
19	0.429	0.008

Gambar 3. Hasil Tabel Melalui Software Rapid Miner

Dari gambar diatas dapat dilihat bahwa terdapat 19 item data yang didapat selama 7 hari, dan dimulai pada tanggal 22-27 Januari 2020. Setelah data diolah ke aplikasi Rapidminer bahwa dapat membantu waitres apabila ada konsumen yang bertanya tentang makanan atau minuman paling Favorit sesuai kebutuhan konsumen dapat mempermudah konsumen.

4. SIMPULAN

Berdasarkan hasil perhitungan secara keseluruhan dari data sampel penjualan diperoleh 52 rule asosiasi yang memenuhi support diatas 20% dan confidence 50%. Berdasarkan hasil pengujian dengan menggunakan Rapidminer dapat disimpulkan bahwa produk yang lebih banya diminati oleh konsumen adalah minimum confidence diatas 50%.

DAFTAR PUSTAKA

- [1] E. S. Sihombing, A. S. Honggowibowo, and D. Nugraheny, "Implementasi Data Mining Menggunakan Metode Apriori Pada Transaksi Penjualan Barang (Studi Kasus Di Chorus Minimarket)," pp. 17–30.
- [2] N. Agustiani, D. Suhendro, K. Apriori, and A. Pembahasan, "Penerapan Data Mining Metode Apriori Dalam Implementasi Penjualan Di Alfamart," vol. 2, pp. 300–304, 2020.
- [3] D. R. S. P, Agus Perdana Windarto, D. Hartama, and I. S. Damanik, "Penerapan klasifikasi c4.5 dalam meningkatkan sistem pembelajaran mahasiswa," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, pp. 593–597, 2020.
- [4] D. R. S. P, A. P. Windarto, D. Hartama, and E. Irawan, "Analisis Datamining Pada Pengelompokkan Penduduk Yang Menjadi Korban Kekerasan Fisik Menurut Wilayah," *Semin. Nas. Sains Teknol. Inf. SENSASI 2019 ISBN*, pp. 350–356, 2019.
- [5] K. F. Irnanda, A. P. Windarto, I. S. Damanik, and I. Gunawan, "Penerapan K-Means pada Proporsi Individu dengan Keterampilan (Teknologi Informasi dan Komunikasi) TIK Menurut Wilayah," *Semin. Nas. Sains Teknol. Inf.*, no. c, pp. 452–456, 2019.
- [6] K. F. Irnanda and A. P. Windarto, "Penerapan Klasifikasi C4.5 Dalam Meningkatkan Kecakapan Berbahasa Inggris dalam Masyarakat," *Semin. Nas. Teknol. Komput. Sains*, pp. 304–308, 2020.
- [7] C. Astria, D. Hartama, A. P. Windarto, and I. Sudahri, "Pengembangan Metode Datamining K-Medoid Pada Kasus Distribusi Listrik di Indonesia," pp. 276–281, 2020.
- [8] J. R. Informatika, "Analisis Penerapan Asosiasi Untuk Menentukan Transaksi Penjualan Pada What ' S Up Café Dengan Metode Algoritma," vol. 1, no. 4, 2019.
- [9] G. Abdurrahman, "Analisis Aturan Asosiasi Data Transaksi Supermarket Menggunakan Algoritma Apriori," pp. 100–111.