

PENERAPAN DATA MINING DALAM MENGANALISIS POLA PEMBELIAN MENGGUNAKAN ALGORITMA APRIORI

Nurul Hidayati¹, Wiga Maulana Baihaqi²

¹Jurusan Teknologi Informasi, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto, Purwokerto, Indonesia
¹nurulhid371@gmail.com, ²wiga@amikompurwokerto.ac.id

Info Artikel

Sejarah artikel:

Diterima 05 Januari 2024

Revisi 15 Maret 2024

Diterima 12 Agustus 2024

Kata kunci:

Pembelian, Algoritma Apriori, Data Mining

ABSTRAK

Meningkatnya jumlah transaksi saat ini menciptakan kebutuhan analisis yang efektif untuk menggali informasi berharga dari pola yang tersembunyi. Data transaksi dari penjualan yang dilakukan setiap hari menyebabkan penyimpanan data yang besar. Pada dasarnya kumpulan data memiliki informasi yang dapat dimanfaatkan dengan baik.

Berdasarkan permasalahan tersebut diperlukan adanya sistem guna mendapatkan informasi yang menunjang dan meningkatkan pola bisnis menggunakan algoritma apriori. Dengan digunakannya pendekatan apriori diharapkan akan memberikan gambaran hubungan antar barang yang dibeli secara bersamaan. Hasil yang diharapkan berupa laporan hasil data mining pola pembelian pelanggan yang sering membeli produk dalam waktu bersamaan guna menyusun strategi pemasaran.

Penulis yang sesuai:

Nurul Hidayati

Program Studi Teknologi Informasi Fakultas Ilmu Komputer Universitas Amikom Purwokerto

Email: nurulhid371@gmail.com

1. PENDAHULUAN

Perkembangan teknologi saat ini telah membuka peluang dalam analisis data guna mendukung pengambilan keputusan terutama dalam sektor bisnis. Salah satunya adalah menganalisis pola pembelian untuk memahami perilaku pelanggan guna meningkatkan strategi pemasaran. Data transaksi penjualan yang terus bertambah setiap waktu dapat menyebabkan penyimpanan data yang cukup besar. Dari data tersebut kurang dimanfaatkan dengan baik, padahal didalamnya memiliki cukup banyak informasi maupun pola yang bermanfaat.

Persaingan dalam dunia usaha menuntut para pengusaha untuk memikirkan cara dan sarana untuk meningkatkan perdagangan produk yang dijual. Salah satunya adalah dalam bidang penjualan furniture. Meskipun ketersediaan inventaris perabotan merupakan faktor penting, mengetahui produk mana yang biasanya dibeli pelanggan juga sama pentingnya. Status inventaris produk harus selalu dipantau oleh bagian manajemen dan staf yang bertugas sebagai admin yang bertugas. Jika ketersediaan produk tidak terpantau

dengan baik maka timbul permasalahan seperti pelanggan mencari produk di toko lain dan rendahnya nilai stok di persediaan produk.

Berdasarkan permasalahan yang terjadi di atas, maka perlu adanya sebuah aplikasi yang menunjang pengelompokan data sesuai jenisnya yang cenderung muncul secara bersamaan dalam sebuah transaksi. Dalam penelitian ini penulis menerapkan data mining dengan algoritma apriori dan berharap dapat menganalisis pola pembelian pelanggan dengan cara menemukan produk yang sering dibeli dalam satu transaksi secara bersamaan. Algoritma apriori data mining juga dapat digunakan untuk mencari hubungan antar elemen dalam catatan transaksi.

Dengan menggunakan algoritma ini, penulis berharap dapat mengidentifikasi pola pembelian yang sering tumpang tindih sehingga dapat menghasilkan informasi yang berharga untuk pengembangan strategi penjualan. Pentingnya menganalisis perilaku pelanggan telah diakui di berbagai industri di beberapa perusahaan dapat meningkatkan operasional, meningkatkan kepuasan pelanggan dan mengembangkan strategi pemasaran yang lebih tepat sasaran.

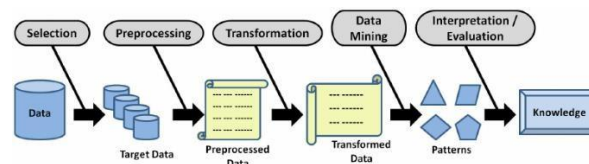
2. METODE

2.1 Pengumpulan Data

Dalam penelitian ini menggunakan dataset yang diperoleh dari web resmi kaggle yang dimana didalamnya terdapat faktor data yang dibutuhkan seperti Order Id, Produk dan Quantiti. Faktor tersebut merupakan data yang digunakan untuk mendapatkan hasil dari penelitian ini.

2.2 Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) merupakan suatu proses mengidentifikasi pola, tren dan informasi penting dalam sebuah data dengan jumlah besar. KDD melibatkan hasil dari proses Data Mining yang kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami.



Gambar 1. Proses Knowledge Discovery in Database (KDD)

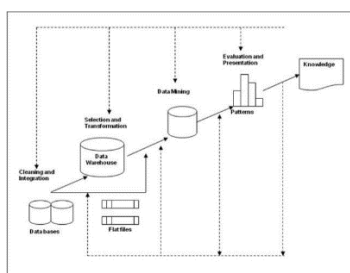
2.3 Data Mining

Pada Data Mining terdapat beberapa tahap yang perlu dilakukan seperti

- a. Data Cleansing
Merupakan proses membersihkan space kosong dan data yang tidak relevan dan data tidak konsisten.
- b. Data Selection
Merupakan proses pemilihan data yang akan digunakan, karena tidak semua data yang ada akan digunakan untuk analisis.
- c. Data Transformation
Merupakan proses perubahan atau penggabungan data ke dalam format yang sesuai.
- d. Proses Mining

Ini merupakan proses utama penerapan metode untuk menentukan informasi berharga yang tersembunyi dalam data.

- e. Pattern Evaluation
Mengidentifikasi pola-pola menarik ke dalam knowledge based.
- f. Evaluation and Presentation
Menjelaskan pengetahuan tentang bagaimana dapat memperoleh informasi yang diperoleh oleh pengguna. Dalam presentasi, visualisasi juga membantu mengkomunikasikan data mining.



Gambar 2. Tahap Data Mining

2.4 Algoritma Apriori

Algoritma Apriori merupakan algoritma yang menggunakan teknik aturan asosiasi guna mencari frekuensi itemset. Dua tolak ukur yang dapat digunakan untuk menentukan apakah hubungan itu penting adalah dukungan dan kepercayaan. *Support* adalah nilai dukungan atau persentase kombinasi elemen-elemen dalam database sementara *confidence* adalah nilai kepastian, kekuatan hubungan antar elemen dalam aturan asosiasi. Proses yang dilakukan pada algoritma apriori untuk meningkatkan frekuensi *itemset* antara lain dengan penggabungan. Proses ini dilakukan dengan menggabungkan elemen dengan elemen lainnya hingga tidak ada kombinasi yang dapat dilakukan. Lalu selanjutnya pemangkasan, yaitu pemangkasan elemen yang telah dikombinasikan menggunakan minimum support yang telah ditentukan.

Gambar 3. Hasil Akhir Menggunakan Persamaan 1

$$Support(A) = \frac{\text{Jumlah Transaksi Mengandung A}}{\text{Total Transaksi}} \times 100$$

3. HASIL DAN PEMBAHASAN

3.1 Ubah Data Transaksi Dalam Bentuk Tabular

Data penjualan selama ini tidak tersusun dengan baik, sehingga data penjualan selama ini tidak dimanfaatkan dengan baik oleh perusahaan. Data tersebut hanya disimpan sebagai arsip maupun pembukuan dalam perusahaan dan tidak diketahui manfaat dari data tersebut. Maka dari itu penulis ambil beberapa transaksi untuk diuji sebagai berikut pada Tabel 1.

Transaksi	Item yang dibeli
1	Office Supplies
2	Furniture
3	Office Supplies, Furniture
4	Office Supplies, Furniture, Technology
5	Office Supplies
6	Office Supplies
7	Office Supplies
8	Office Supplies, Technology
9	Office Supplies
10	Furniture

Tabel 1. Tabel Transaksi

Dari data Tabel 1 penulis pisahkan per item yang dibeli pada masing masing transaksi kedalam bentuk tabular yang tertulis pada Tabel 2

Transaksi	Item yang dibeli		
	Office Supplies	Furniture	Technology
1	1	0	0
2	0	1	0
3	1	1	0
4	1	1	1
5	1	0	0
6	1	0	0
7	1	0	0
8	1	0	1
9	1	0	0
10	0	1	0
Jumlah	8	4	2

Tabel 2. Bentuk Tabular Data Transaksi

3.2 Melakukan Kombinasi 2 Calon Itemset

Mengkombinasikan 2 calon *itemset* pada setiap transaksi dan frekuensi dihitung sesuai data tabular pada Tabel 3.

Kombinasi	Jumlah
Office Supplies, Furniture	2
Office Supplies, Technology	2
Furniture, Technology	1

Tabel 3. Calon 2 itemset**3.3 Menentukan Minimal Nilai Support**

Berdasarkan data pada Tabel 3 ditetapkan bahwa minimum nilai supportnya = 1. Maka $F2 = \{\{\text{Office Supplies, Furniture}\}, \{\text{Office Supplies, Technology}\}, \{\text{Furniture, Technology}\}\}$

3.4 Melakukan Kombinasi 3 Calon Itemset

Kombinasi 3 *itemset* didapatkan dari $F2$. Hanya ada 1 calon 3 *itemset* dikarenakan hanya memiliki 3 item produk saja.

Kombinasi	Jumlah
Office Supplies, Furniture, Technology	1

Tabel 4. Calon 2 itemset

Dengan demikian $F3 = \{\text{Office Supplies, Furniture, Technology}\}$

3.5 Pembentukan Aturan Asosiasi

Pembentukan aturan asosiasi dapat dicari setelah semua pola frekuensi tinggi ditentukan.

Aturan	Confidence	
Jika membeli Office Supplies dan Furniture, maka akan membeli Technology	1/2	50%
Jika membeli Office Supplies dan Technology, maka akan membeli Furniture	1/2	50%
Jika membeli Technology dan Furniture, maka akan membeli Office Supplies	1/1	100%

Tabel 5. Aturan Asosiasi 3 *itemset*

Misalkan ditetapkan nilai minimum confidence 50%, maka aturan yang dapat terbentuk

- Jika membeli Office Supplies dan Furniture, maka akan membeli Technology.
- Jika membeli Office Supplies dan Technology, maka akan membeli Furniture.
- Jika membeli Technology dan Furniture, maka akan membeli Office Supplies.

Aturan asosiasi dari $F2$ dapat dilihat pada Tabel 6 berikut

Aturan	Confidence	
Jika membeli Office Supplies, maka akan membeli Furniture	2/8	25%
Jika membeli Furniture, maka akan membeli Office Supplies	2/4	50%
Jika membeli Office Supplies, maka akan membeli Technology	2/8	25%
Jika membeli Technology, maka akan membeli Office Supplies	2/2	100%
Jika membeli Furniture, maka akan membeli Technology	1/4	25%
Jika membeli Technology, maka akan membeli Furniture	1/2	50%

Tabel 6. Aturan Asosiasi 2 *itemset*

Aturan Asosiasi Final berdasarkan *support x confidence* terbesar dapat dilihat pada Tabel 7

Aturan	Support	Confidence	Support X Confidence
Jika membeli Furniture, maka akan membeli Office Supplies	20%	50%	10%
Jika membeli Technology, maka akan membeli Office Supplies	20%	100%	20%
Jika membeli Technology, maka akan membeli Furniture	10%	50%	5%
Jika membeli Office Supplies dan Furniture, maka akan membeli Technology	10%	50%	5%
Jika membeli Technology dan Furniture, maka akan membeli Office Supplies	10%	100%	10%

Tabel 7. Aturan Final Asosiasi

Tabel *final association rule* menggambarkan *support* dan *confidence* untuk setiap kombinasi dua item dan tiga item. Tabel *final association rule* mendukung hasil perhitungan jumlah transaksi yang mengandung A dan B dibagi dengan total transaksi. Sebaliknya, *confidence* dihitung dengan membagi jumlah transaksi yang mengandung A dan B dengan jumlah transaksi yang mengandung A. Produk *support* dan *confidence* merupakan hasil akhir dari algoritma apriori.

4. KESIMPULAN

Dari hasil analisa dan pembahasan dapat disimpulkan bahwa data mining sangat berguna untuk mengetahui hubungan pola frekuensi penjualan produk yang paling sering dibeli saat melakukan transaksi. Digunakannya algoritma apriori untuk menyusun strategi pemasaran dalam memasarkan produk dengan meneliti pola pembelian dapat disimpulkan bahwa nama produk yang paling banyak terjual dengan minimum support 10% adalah Office Supplies, Furniture, Technology. Penelitian ini menghasilkan penerapan teknik Data Mining menggunakan Algoritma Apriori dalam menganalisis pola pembelian pelanggan. Dengan menggali transaksi pembelian, penulis berhasil mengidentifikasi pola pola yang signifikan dalam perilaku pembelian pelanggan. Beberapa kesimpulan utama yang dapat diambil dari penelitian ini yaitu konsumen pada saat membeli beberapa produk sekaligus sebagai alat alternatif pendukung keputusan dalam menemukan kecenderungan pola kombinasi kelompok barang dan menentukan penempatan produk pada area yang berdekatan sesuai dengan perilaku konsumen. Dapat digunakan sebagai informasi untuk menentukan kebutuhan masyarakat sebagai tindakan alternatif untuk membantu mengidentifikasi produk yang jarang dibeli konsumen pada saat membeli produk pada waktu yang bersamaan dan untuk meningkatkan strategi pemasaran dengan menawarkan diskon pada produk tertentu terutama yang jarang dibeli untuk menarik minat pembelian konsumen.

REFERENSI

- [1] Yanto, Robi, and Riri Khoiriah. "Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat." *Creative Information Technology Journal* 2.2 (2015): 102-113.
- [2] Pramudiono, I., 2007, Pengantar Data Mining: Menambang Permata Pengetahuan di Gunung Data. diakses tanggal 15 mei 2015.
- [3] Han, J and Kamber, M, 2006 " Data Mining Concepts and Techniques, Second Edition". Morgan Kaufman, San Francisco. <http://id.wikipedia.org/wiki/Elektronika>, 23 Maret 2015.
- [4] Erwin, 2009. Analisis Market Basket dengan Algoritma Apriori dan FP-Growth. *Jurnal Generik* Vol.4 No2. juli 2009.
- [5] M. P. Tana, F. Marisa, and I. D. Wijaya, "Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Pada Toko Oase Menggunakan Algoritma Apriori," *J I M P - J. Inform. Merdeka Pasuruan*, vol. 3, no. 2, Aug. 2018.



- [6] M. Safii and A. Trydillah, "Implementasi Data Mining Dalam Menentukan Pola Pembelian," vol. 3, no. 1, pp. 66–71, 2019.
- [7] M. Fauzy, K. R. S. W, and I. Asror, "PENERAPAN METODE ASSOCIATION RULE MENGGUNAKAN ALGORITMA APRIORI PADA SIMULASI PREDIKSI HUJAN WILAYAH KOTA BANDUNG," *Utama*, vol. 2, no. 2, 2016.
- [8] Sugiyantno and A. P. Widyasari, "APLIKASI MONITORING KETERSEDIAAN STOK BARANG MINIMARKET DENGAN METODE MARKET BASKET ANALYSIS (MBA)," in *Seminar Nasional Teknologi Informasi dan Multimedia 2016*, 2016, pp. 6–7.
- [9] N. F. Fahrudin, "Penerapan Algoritma Apriori untuk Market Basket Analysis," *MINN J.*, vol. 4, no. 1, 2019.
- [10] Santoso,B.,2007, *Data Mining :Teknik pemanfaatan data untuk keperluan bisnis*, Graha Ilmu, Yogyakarta.
- [11] H. Hutahaean, B. Sinaga, and A. Rajagukguk, "ANALISA DAN PERANCANGAN APLIKASI ALGORITMA APRIORI UNTUK KORELASI PENJUALAN PRODUK (STUDI KASUS : APOTIK DIORY FARMA)," vol. 1, no. 1, pp. 7–13, 2016.

