

Analisa Data Mining Terhadap Penjualan Food Dengan Metode Apriori Pada Kopsyahira

Agung Riyanto¹, Melan Susanti²

STMIK Nusa Mandiri^{1,2}

agungriyanto688@gmail.com¹, melan.msu@nusamandiri.ac.id²

Abstrak — Setiap perusahaan yang bergerak dalam bidang perdagangan harus mempunyai strategi dalam peningkatan pelayanan. Beberapa diantaranya mengatur susunan barang (display) atau membuat tampilan dari sebuah toko terlihat menarik dan mempermudah belanja sehingga konsumen bersedia datang kembali untuk berbelanja. Banyak transaksi disetiap harinya namun masih dilakukan dengan cara manual. Sehingga mungkin saja terjadi banyak kesalahan serta tidak akuratnya laporan. Banyaknya transaksi juga hanya dijadikan dokumen saja. Tidak menutup kemungkinan data transaksi tersebut banyak yang hilang atau terselip. Kumpulan data transaksi jika dibiarkan saja selama berbulan-bulan, maka data tersebut hanya akan menjadi data yang tidak berarti dan akan menjadi faktor penghambat dalam peningkatan pelayanan. Pembelian sering sekali dilakukan secara bersamaan dalam satu waktu, sehingga terjadi antrian didalam toko. Di Kopsyahira juga terjadi beberapa kendala dalam hal penjualan, terutama penjualan food. Dalam penelitian ini, peneliti akan memanfaatkan algoritma Apriori, penulis menggunakan software data mining Tanagra. Hasil dari penelitian ini menghasilkan 2 aturan asosiasi final jika menggunakan minimum support 30% dan Confidence 66%.

Kata Kunci: Data Mining, Apriori, Penjualan food.

Abstract - Every company engaged in trade must have a strategy to improve service. Some of them regulate the arrangement of goods (display) or make the appearance of a store look attractive and make shopping easier so that consumers are willing to come back to shop. Many transactions every day but are still done manually. So there might be a lot of errors and inaccurate reports. The number of transactions is also only used as a document. It is not possible for many transaction data to be lost or tucked away. Collection of transaction data if left alone for months, then the data will only be meaningless data and will be a limiting factor in improving services. Purchases are often done simultaneously at one time, so there is a queue in the store. In Kopsyahira there were also several obstacles in terms of sales, especially food sales. In this study, researchers will use the Apriori algorithm, the author uses Tanagra's data mining software. The results of this study produce 2 final association rules if using a minimum support of 30% and Confidence of 66%.

Keywords: Data Mining, Apriori, Food Sales

I. PENDAHULUAN

Koperasi Syariah Ikhlas Sejahtera (Kopsyahira) banyak menjual produk. Banyak transaksi disetiap harinya namun masih dilakukan dengan cara manual. Sehingga mungkin saja terjadi banyak kesalahan serta tidak akuratnya laporan. Banyaknya transaksi juga hanya dijadikan dokumen saja. Tidak menutup kemungkinan data transaksi tersebut banyak yang hilang atau terselip. Kumpulan data transaksi jika dibiarkan saja selama berbulan-bulan, maka data tersebut hanya akan menjadi data yang tidak berarti dan akan menjadi faktor penghambat dalam peningkatan pelayanan. Pembelian sering sekali dilakukan secara bersamaan dalam satu waktu, sehingga terjadi antrian didalam toko. Oleh Karena itu, pada penelitian ini penulis mengelompokan produk yang tersedia menjadi dua katagori, yaitu food dan non food. Untuk katagori food meliputi produk makanan dan minuman, sedangkan produk non food meliputi produk rumah tangga, ATK, alat perlengkapan. Penempatan barang dagangan yang tidak berpola kadang membuat

situasi menjadi tidak teratur. Konsumen harus berputar putar mencari barang yang diinginkan karena kondisi tersebut. Sehingga membutuhkan waktu yang cukup lama untuk berbelanja.

Semakin banyak transaksi penjualan setiap harinya maka data penjualan pun semakin lama akan bertambah semakin banyak. Jika dibiarkan saja, maka data-data transaksi penjualan tersebut hanya menjadi sampah yang tidak berarti. Dengan adanya dukungan perkembangan teknologi, semakin berkembang pula kemampuan dalam mengumpulkan dan mengolah data. Pemanfaatan informasi dan pengetahuan yang terkandung di dalam banyaknya data tersebut, pada saat ini disebut dengan data mining. Data-data transaksi tersebut tidak hanya berfungsi sebagai arsip bagi perusahaan, data tersebut dapat dimanfaatkan dan diolah menjadi informasi yang berguna untuk peningkatan penjualan dan promosi produk (Purnia & Warnilah, 2017).

Dengan demikian, penulis memandang perlunya sistem yang dapat mendukung perusahaan dalam mengambil keputusan

secara cepat, tepat, serta bisa meningkatkan penjualan. Oleh karena itu dibutuhkan sistem pengolahan database menggunakan data mining dengan metode Algoritma Apriori. Dengan metode ini diharapkan bisa membantu Koperasi Syariah Ikhlas Sejahtera (Kopsyahira) mampu menemukan pola penjualan produk yang saling berasosiasi. Sehingga perusahaan bisa meningkatkan jumlah stok produk-produk yang saling berasosiasi tersebut.

1. Tinjauan Pustaka

Teori-teori atau tinjauan pustakan yang berkaitan dengan penelitian ini yaitu:

1.1. Data Mining

Data mining adalah langkah analisis terhadap proses penemuan pengetahuan di dalam basisdata atau *Knowledge Discovery in Database* yang disingkat KDD (Suyanto, 2017). Data Mining adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika (Husin & Mulyaningsih, 2015). Pada tahap *Knowledge Discovery Database* (KDD) terdapat beberapa fase yaitu sebagai berikut:

1. Seleksi Data (*Selection*)

Selection (Seleksi/Pemilihan) data dari merupakan sekumpulan data operasional perlu dilakukan sebelum tahap pengalihan informasi dalam *Knowledge Discovery Database* (KDD) dimulai. Data Hasil seleksi yang akan digunakan untuk proses data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional

2. Pemilihan Data (*Preprocessing/Cleaning*)

Pada Proses ini mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses enrichment, yaitu proses "memperkaya" data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

3. Transformasi (*Transformation*)

Pada fase ini yang dilakukan adalah mentransformasi bentuk data yang belum memiliki entitas yang jelas ke dalam bentuk data yang valid atau siap untuk dilakukan proses data mining.

4. Data Mining

Pada tahap ini yang dilakukan adalah menerapkan algoritma atau metode pencarian pengetahuan.

5. Interpretasi/Evaluasi

Pada tahap terakhir ini yang dilakukannya adalah proses pembentukan keluaran yang mudah dimengerti yang bersumber pada proses data mining pola informasi.

1.2. Algoritma Apriori

Algoritma Apriori adalah algoritma pengambilan data dengan aturan asosiatif (*Association rule*) untuk menentukan hubungan asosiatif suatu kombinasi item. Ada dua pengukuran penting untuk sebuah rule yaitu *support* dan *confidence*. *Support* adalah suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu item atau itemset dari keseluruhan transaksi, sedangkan *Confidence* adalah suatu ukuran yang menunjukkan hubungan antar dua item secara kondisional (Kurniawati, Kusuma, & Dewansyah, 2019). Untuk menghitung support 1 item set bisa menggunakan persamaan sebagai berikut:

$$Support(A) = \frac{\Sigma \text{Transaksi yang mengandung } A}{\Sigma \text{Transaksi}} \quad (1)$$

Untuk menghitung support 2 itemset dapat diselesaikan dengan persamaan berikut :

$$Support(A,B) = P(A \cap B) = \frac{\Sigma \text{Transaksi yang mengandung } A \text{ dan } B}{\Sigma \text{Transaksi}} \quad (2)$$

Untuk menghitung support 3 item set dapat diselesaikan dengan rumus berikut :

$$Support(A,B,C) = \frac{\Sigma \text{Transaksi yang mengandung } A, B \text{ dan } C}{\Sigma \text{Transaksi}} \quad (3)$$

Untuk menghitung *Confidence* dapat diselesaikan dengan rumus :

$$Confidence = P(A/B) = \frac{\Sigma \text{Transaksi mengandung } A \text{ dan } B}{\Sigma \text{Transaksi } A} \quad (4)$$

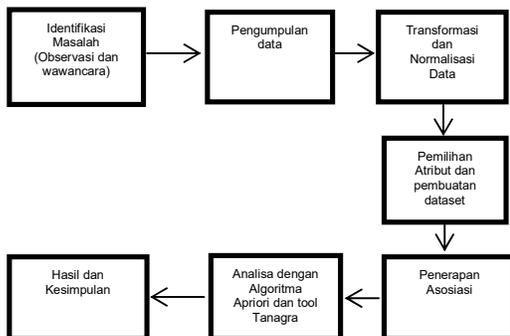
1.3. Tanagra

Menurut Julsam, Tanagra adalah *software* Data Mining bebas untuk tujuan akademik dan penelitian ini mengusulkan beberapa metode data mining dari analisis eksplorasi data, pembelajaran statistik, pembelajaran mesin dan daerah database (Badrul, 2016). Tanagra adalah "proyek open source" karena setiap peneliti dapat mengakses ke kode sumber, dan menambahkan algoritma sendiri, sejauh dia setuju dan sesuai dengan lisensi distribusi perangkat lunak. Tujuan utama dari proyek Tanagra adalah memberikan peneliti dan mahasiswa yang mudah untuk menggunakan perangkat lunak data mining, sesuai dengan norma-norma yang hadir dari pengembangan perangkat lunak dalam domain ini (terutama dalam desain GUI dan cara untuk menggunakannya), dan memungkinkan untuk menganalisis baik data yang nyata atau sintesis. Tujuan kedua Tanagra adalah untuk

mengusulkan kepada peneliti arsitektur yang memungkinkan mereka untuk dengan mudah menambahkan metode penambangan datamereka sendiri, untuk membandingkan kinerja mereka.

II. METODOLOGI PENELITIAN

Penelitian adalah suatu proses, yaitu rangkaian langkah-langkah yang dilakukan secara terencana dan sistematis guna mendapatkan pemecahan masalah atau mendapatkan jawaban terhadap pertanyaan-pertanyaan tertentu (Suryabrata, 2014). Langkah-langkah penelitian diawali dengan identifikasi masalah dengan cara melakukan observasi dan wawancara di KOPSYAHIRA (Koperasi Syariah Ikhlas Sejahtera). Setelah itu penulis mengumpulkan data penjualan food yang nantinya akan penulis analisa. Data yang telah terkumpul maka akan dilakukan transformasi dan normalisasi agar tidak terjadi data yang berulang. Lalu akan dilakukan pemilihan atribut dan pembuatan dataset. Setelah itu aturan Apriori akan ditetapkan serta akan dibantu menggunakan tool Tanagra untuk dianalisa sehingga mendapatkan hasil pengujian dan kesimpulan.



Sumber: Penelitian (2019)

Gambar 1. Tahapan Peneitian

III. HASIL DAN PEMBAHASAN

Tahap dalam menganalisa data dengan algoritma apriori pada penjualan produk food dimulai dengan menyeleksi dan membersihkan data - data yang akan dianalisis, kemudian dicari semua jenis item produk food yang ada didalam transaksi penjualan. Selanjutnya dicari jumlah setiap item yang ada pada transaksi penjualan. Pada penelitian ini mengambil seample 25 produk food yang paling banyak terjual selama 6 bulan, dimulai bulan agustus 2018 sampai bulan desember 2018. Selanjutnya untuk mempermudah penelitian dibuat kode untuk 25 produk food tersebut yang dapat dilihat pada tabel 1:

Tabel 1. Produk Food

No	Kode Food	Nama Minuman
1	F-032	Roti Animo
2	F-033	Frestea Imut
3	F-034	Teh Pucuk
4	F-035	Teh Kotak
5	F-036	Coca-Cola 250 MI Seru
6	F-037	Pocari Kaleng
7	F-038	Susu Bendera
8	F-039	Pocari Botol 350
9	F-040	Good Day
10	F-041	Milo Kotak
11	F-042	Mizone
12	F-043	Nescafe Kotak
13	F-044	Frestea Botol
14	F-045	Keripik Pisang
15	F-046	Yakult
16	F-047	Kacang Kapri
17	F-048	Ultra Kotak 200ml
18	F-049	Sprite 425 MI Seru
19	F-050	Sprite Botol
20	F-051	White Cofie
21	F-052	Indocafe Capucino
22	F-053	Fanta 250 MI Seru
23	F-054	Nescaffe Kaleng
24	F-055	Pulpy Orange Botol
25	F-056	Ades Kecil Botol 300ml

Sumber: Kopsyahira (2018)

1. Analisa Pola Frekuensi Tinggi

Proses pembentukan C1 atau disebut dengan 1 itemset dengan jumlah minimum support = 30%. Berikut merupakan perhitungan pembentukan 1 itemset :

$$S(F-032) = \frac{\sum F-032}{\sum \text{Transaksi}} = \frac{63}{120} * 100\% = 52,50\%$$

$$S(F-033) = \frac{\sum F-033}{\sum \text{Transaksi}} = \frac{26}{120} * 100\% = 21,67\%$$

Hasil perhitungan minimum support untuk 1 itemset pada 25 produk food dapat dilihat pada tabel 2:

Tabel 2. Hasil Perhitungan 1 Itemset

Itemset	Jumlah	Support
F-032	63	52,50 %
F-033	26	21,67 %
F-034	40	33,33 %
F-035	49	40,83 %
F-036	17	14,17 %
F-037	11	9,17 %
F-038	14	11,67 %
F-039	11	9,17 %
F-040	16	13,33 %
F-041	10	8,33%
F-042	11	9,17 %
F-043	12	10,00 %
F-044	48	40,00 %
F-045	45	37,50 %
F-046	6	5,00 %
F-047	15	12,50 %
F-048	8	6,67 %
F-049	10	8,33 %

F-050	10	8,33 %
F-051	6	5,00 %
F-052	7	5,83 %
F-053	11	9,17 %
F-054	11	9,17 %
F-055	14	11,67 %

Sumber: Penelitian (2019)

Dari hasil perhitungan 1 itemset yang memenuhi minimum support 30 % dapat dilihat pada tabel 3:

Tabel 3. Hasil Perhitungan 1 Itemset Memenuhi Minimal Support

Itemset	Jumlah	Support
F-032	63	52,50 %
F-034	40	33,33 %
F-035	49	40,83 %
F-044	48	40,00 %
F-045	45	37,50 %
F-056	40	33,33 %

Sumber: Penelitian (2019)

Proses pembentukan C2 atau disebut dengan 2 itemset dengan perhitungan sebagai berikut:

$$S(F-032, F-034) = \frac{\Sigma \text{Transaksi } F-032, F-034}{\Sigma \text{Transaksi}} = \frac{21}{120} * 100\% = 17,50\%$$

$$S(F-032, F-035) = \frac{\Sigma \text{Transaksi } F-032, F-035}{\Sigma \text{Transaksi}} = \frac{37}{120} * 100\% = 30,83\%$$

Hasil perhitungan minimum support untuk 2 itemset dapat dilihat pada tabel 4:

Tabel 4. Hasil Perhitungan 2 Itemset

Itemset	Jumlah	Support
F-032, F-034	21	17,50 %
F-032, F-035	37	30,83 %
F-032, F-044	38	31,67 %
F-032, F-045	21	17,50 %
F-032, F-056	5	4,17 %
F-034, F-035	17	14,17 %
F-034, F-044	15	12,50 %
F-034, F-045	20	16,67 %
F-034, F-056	6	5,00 %
F-035, F-044	32	26,67 %
F-035, F-045	13	10,83 %
F-035, F-056	3	2,50 %
F-044, F-045	12	10,00 %
F-044, F-056	4	3,33 %
F-045, F-056	19	15,83 %

Sumber: Penelitian (2019)

Dari hasil perhitungan 2 itemset yang memenuhi minimum support 30 % dapat dilihat pada tabel 5:

Tabel 5. Hasil Perhitungan 2 Itemset Memenuhi Minimal Support

Itemset	Jumlah	Support
F-032, F-044	38	31,67 %
F-032, F-035	37	30,83 %

Sumber: Penelitian (2019)

2. Pembentukan Aturan Asosiasi

Setelah semua pola *frekuensi* tinggi ditemukan, baru dicari aturan *asosiasi* yang memenuhi syarat minimum *Confidence* dengan menghitung *Confidence* aturan asosiatif A→B.

Minimal *Confidence* = 66%

$$Confidence = P(F-032, F-035) = \frac{\Sigma \text{Transaksi } F-032, F-035}{\Sigma F-032} = \frac{37}{63} * 100\% = 58,73\%$$

Confidence = P(F-035, F-032)

$$= \frac{\Sigma \text{Transaksi } F-035, F-032}{\Sigma F-035} = \frac{37}{49} * 100\% = 75,51\%$$

Dari kombinasi 2 itemset yang telah ditemukan, dapat dilihat besarnya nilai support, dan confidence dari calon aturan asosiasi seperti tampak pada tabel 6:

Tabel 6. Hasil Perhitungan Nilai Confidence

Aturan	Support	Confidence
Jika membeli Roti Animo (F-032), maka akan membeli Teh Kotak (F-035)	37/63	58,73%
Jika membeli Teh Kotak (F-035), maka akan membeli Roti Animo (F-032)	37/49	75,51 %
Jika membeli Roti Animo (F-032), maka akan membeli Frestea Botol (F-044)	38/63	60,32 %
Jika membeli Frestea Botol (F-044), maka akan membeli Roti Animo (F-032)	38/48	79,17 %

Sumber: Penelitian (2019)

3. Aturan Asosiasi Final

Berdasarkan dari calon aturan asosiasi pada Tabel 6 maka yang memenuhi minimal support 30% dan minimal confidence 66% dapat dilihat pada tabel 7:

Tabel 7. Hasil 2 itemset yang memenuhi Minimal Confidence

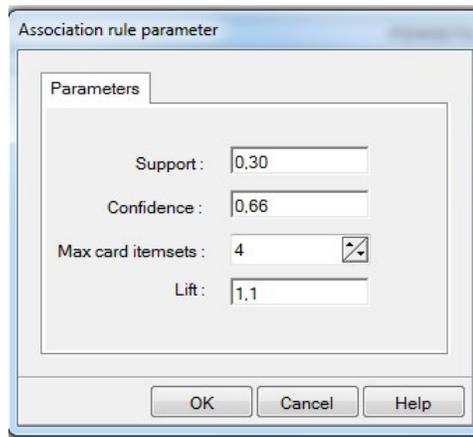
Aturan	Support	Confidence
Jika membeli Frestea Botol (F-044), maka akan membeli Roti Animo (F-032)	31,67 %	79,17 %
Jika membeli Teh Kotak (F-035), maka akan membeli Roti Animo (F-032)	30,83 %	75,51 %

Sumber: Penelitian (2019)

Berdasarkan tabel 7, produk food yang paling sering dibeli secara bersamaan adalah Roti Animo, Frestea Botol, Teh Kotak. Dengan diketahui produk yang paling dibeli secara bersamaan tersebut, sehingga perusahaan dapat meneliti apa kelebihan produk-produk tersebut dengan produk lainnya dan dapat menambah persediaan produk food tersebut di Kopsyahira.

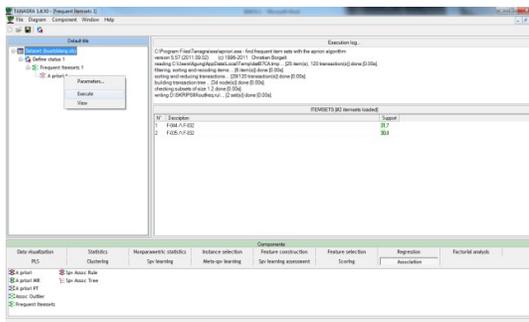
4. Pengolahan Data dengan Tanagra

Pada Tanagra diinputkan parameter untuk nilai minimum support dan nilai confidence (Association Rule) seperti terlihat pada gambar 2 :



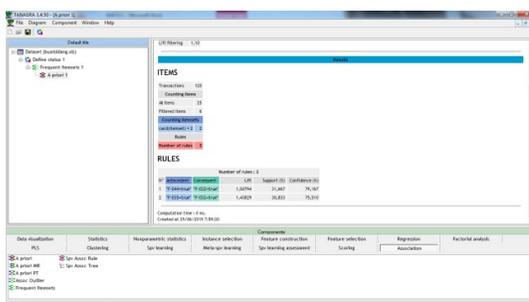
Sumber: Penelitian (2019)
Gambar 2. Tampilan Association Rule

Selanjutnya memproses Asosiasi Final dengan mengklik kanan pada Apriori kemudian klik Execute seperti yang terlihat pada gambar 3:



Sumber: Penelitian (2019)
Gambar 3. Tampilan Proses Asosiasi Final

Menampilkan hasil Asosiasi Final, klik kanan pada Apriori kemudian klik View maka akan tampil hasil Asosiasi Final seperti yang terlihat pada gambar 4:



Sumber: Penelitian (2019)
Gambar 4. Tampilan Hasil Asosiasi Final

Dari hasil yang diproses dengan Tanagra menunjukkan hasil yang sama dengan perhitungan manual.

IV. KESIMPULAN

Berdasarkan hasil pengumpulan dan analisis data dalam penelitian ini, maka dapat diambil kesimpulan bahwa dengan menggunakan metode Apriori dengan menggunakan Tanagra dapat membantu dalam mengetahui produk yang paling banyak terjual, yang pada akhirnya dapat digunakan untuk menyusun strategi penjualan dan dapat dilakukan promosi dengan melakukan kombinasi penjualan antara produk yang laku dengan produk yang kurang laku. Untuk membantu dalam menyusun strategi penjualan selanjutnya dengan menggunakan metode Apriori bisa menentukan produk yang kurang laku dan bisa memberikan solusi untuk meminimalis penjualan produk yang kurang laku dengan mengurangi jumlah pesanan produk tersebut.

V. REFERENSI

- Badrul, M. (2016). Algoritma Asosiasi Dengan Algoritma Apriori Untuk Analisa Data Penjualan. *Jurnal PILAR Nusa Mandiri*, 12(2), 121–129.
- Husin, A. I., & Mulyaningsih, F. (2015). Penerapan Metode Data Mining Analisis Terhadap Data Penjualan Pakaian Dengan Algoritma Apriori. *Snipstek*, 45–56.
- Kurniawati, L., Kusuma, A. E., & Dewansyah, B. (2019). Implementasi Algoritma Apriori Untuk Menentukan Persediaan Spare Part Compressor. *Journal of Computer Engineering System and Science*, 4(1), 6–9.
<https://doi.org/10.24114/cess.v4i1.11303>
- Purnia, D. S., & Warnilah, A. I. (2017). Implementasi Data Mining Pada Penjualan Kacamata Menggunakan Algoritma Apriori. *Indonesian Journal on Computer and Information Technology*, 2(2), 31–39.
- Suryabrata, S. (2014). *Metodologi Penelitian*. Jakarta: PT RajaGrafindo Persada.
- Suyanto. (2017). *Data Mining Untuk Klasifikasi dan Klusterisasi Data*. Bandung: Informatika Bandung.