
PERANCANGAN APLIKASI DATA MINING TRANSAKSI PENJUALAN UNTUK MENGETAHUI POLA BELI KONSUMEN PADA TOKO SINGGALANG PADANG MENGGUNAKAN ALGORITMA APRIORI BERBASIS WEB

Nelisa, Aulia Fitrul Hadi

Universitas Putra Indonesia YPTK Padang

Email : nelisanasution24@gmail.com,fitrulhadi@upiyptk.ac.id

ABSTRACT

It takes a method or technique that can transform mountains of data into a valuable information or knowledge (*knowledge*) that are useful to *support* business decision making. Therefore in this paper association analysis application developed for extracting and interpreting the pattern of trend of sales of goods are often sold simultaneously from the transaction data using *algorithms apriori*.

Apriori algorithms will form a frequent itemset as predetermined based on two parameters, *support* and *confidence*, to find the rules of the *association* between a combination of items. *Knowledge* of a product can be used by companies to increase production and sales of a product.

Keywords: Data Mining, Algorithm Apriori, the Association Rules, Sales.

1. Pendahuluan

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam *data mining* sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan (Jayapana, 2015).

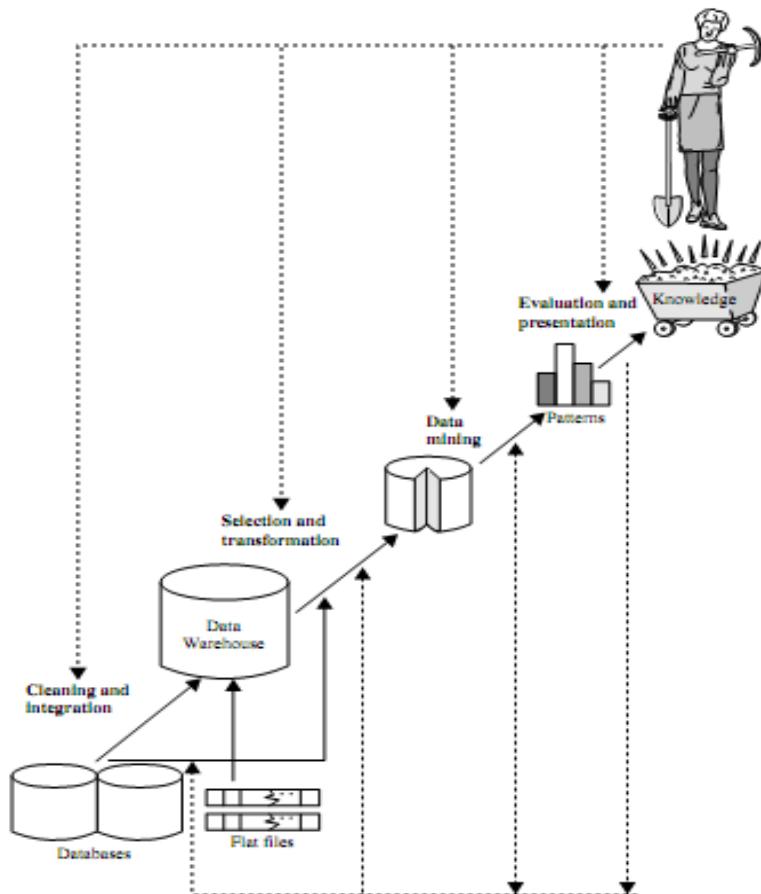
Data penjualan yang sudah ada akan diolah atau dianalisis untuk mengetahui tingkat kecenderungan konsumen di setiap tempat tujuan pemasaran produk pada faktor ketertarikannya. Dari pengolahan data tersebut akan diperoleh suatu pola konsumsi masyarakat terhadap produk dari perusahaan tersebut.

2. Landasan Teori

2.1. Definisi Data Mining

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar. Istilah *data mining* memiliki hakikat sebagai disiplin ilmu yang tujuannya adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki. *Data mining*, sering juga disebut sebagai *Knowledge Discovery in Database* (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar (Jurnal : Mujib Ridwan, dkk 2013).

Sebagai suatu rangkaian proses, *data mining* dapat dibagi menjadi beberapa tahap proses yang diilustrasikan pada Gambar 2.4. Tahap-tahap tersebut bersifat interaktif, pemakai terlibat langsung atau dengan perantaraan *knowledge base* (Jurnal : Mujib Ridwan, dkk 2013).



Gambar 2.4 Tahapan-Tahapan Data Mining

Tahap-tahap *data mining* adalah sebagai berikut:

1. **Pembersihan data (*data cleaning*)**

Pembersihan data merupakan proses menghilangkan *noise* dan data yang tidak konsisten atau data tidak relevan.

2. **Integrasi data (*data integration*)**

Integrasi data merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru.

3. **Seleksi data (*data selection*)**

Data yang ada pada *database* sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*.

4. **Transformasi data (*data transformation*)**

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining*.

5. **Proses *mining***

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

6. **Evaluasi pola (*pattern evaluation*)**

Untuk mengidentifikasi pola-pola menarik kedalam *knowledge based* yang ditemukan.

7. **Presentasi pengetahuan (*knowledge presentation*)**

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna (Jurnal : Mujib Ridwan, dkk 2013).

2.2.Teknik *Data Mining*

Beberapa teknik dan sifat *Data mining*, yaitu (Fajar. 2013:14) :

1. Classification

Klasifikasi (Classification) adalah proses untuk menemukan model atau fungsi yang menjelaskan untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atas kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu obyek yang labelnya tidak diketahui.. (Wahyu Pramusinto, William Budi Utama Gunamawan, 2013 : 2).

2. Clustering

Mempartisi data-set menjadi beberapa sub-set atau kelompok sedemikian rupa sehingga elemen-elemen dari suatu kelompok tertentu memiliki set properti yang di-share bersama, dengan tingkat similaritas yang tinggi dalam suatu kelompok dan tingkat similaritas antar kelompok yang rendah. Disebut juga *unsupervised learning* (Fajar. 2013).

3. Association Rule Discovery

Analisis asosiasi adalah teknik *mining* untuk menemukan aturan asosiatif antara suatu kombinasi item. analysis (Wahyu Pramusinto, William Budi Utama Gunamawan, 2013 : 2).

4. Sequential Pattern Discovery

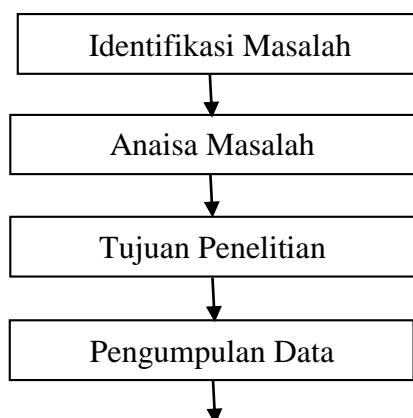
Mencari sejumlah event yang secara umum terjadi bersama-sama.

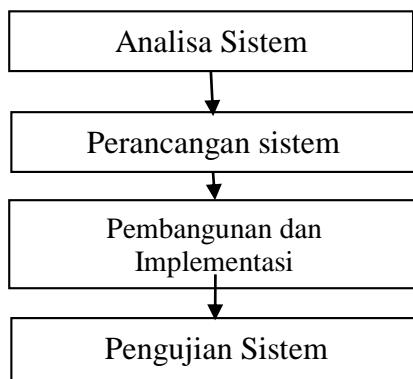
5. Regression

Memprediksi nilai dari suatu variable kontinyu yang diberikan berdasarkan nilai dari variable yang lain, dengan mengansumsikan sebuah model ketergantungan linier atau nonlinier. Teknik ini dipelajari dalam statistika, bidang jaringan saraf tiruan (*neural network*) (Fajar. 2013).

3. Metodologi Penelitian

Kerangka penelitian penelitian ini salah satu metode yang dapat memudahkan seorang peneliti dalam menyelesaikan laporan dan penelitian adalah dengan menyusun sebuah kerangka penelitian, agar penelitian tidak keluar dari objek yang akan di teliti Adapun kerangka penelitian yang penulis lakukan dapat tergambar seperti bagan dibawah ini.





Gambar 3.1 Kerangka kerja Penelitian

4. Analisa dan Pengolahan Data

4.1 Analisa Data Mining dengan Algoritma Apriori

Analisa data merupakan tahap melakukan penganalisaan terhadap data penjualan yang diperlukan untuk perancangan *system*, pada penganalisa dilakukan pengolahan data sesuai dengan ketentuan rumus yang telah ditetapkan.

Dimana dalam proses perhitungannya hanya data transaksi penjualan yang akan diolah dalam algoritma apriori, aturan *asosiasi* dilaksanakan dalam 2 langkah proses, yaitu:

1. Temukan semua *frequency itemset* ; Berdasarkan defenisi, masing – masing dari *itemset* maka muncul sedikitnya dengan *frequency* sebesar diberikan dalam *minimum support count*.
2. Munculkan *strong association rule* dari *frequency itemset*, Berdasarkan defenisi, aturan ini harus memenuhi *minimum support* dan *minimum confidence*.

Adapun data-data yang dipelukan dalam penelitian ini dapat dilihat pada tabel 4.1 berikut:

<i>Id Transaksi</i>	<i>Item yang dibeli</i>
1	Gery Salut, Roma Malkist, Nescafe Original, Aqua , Fruit Tea, Pulpy Orange, Milo, Sari Roti.
2	Milo, Sampurna 16, Mizone, FOX 90G, Bear Brand.
3	Pulpy Orange, Kopiko, Oreo Mini, Astor Wafer, Chunky Bar, Marlboro.
4	Tango Wafer, Taro Net, Djava Young Bakery, Ohayo Bakery, Qtela , Putri Chania Pukis, Koko Krunch Combo.
5	Aqua, Lasegar, Bear Brand, Buavita, Fanta Strawbery.
6	Aqua, Peyek Anugrah, Ohayo Bakery, Teh Pucuk, Aim Biskuit Crispy, Putri Chania Pukis, Tebs Tea Cooler, Sari Roti, MM Nutriboost.
7	Teh Botol, Milo, Nescafe Latte, Mkp Cap Lang.
8	Aqua, Pulpy Orange, Kopiko, Astor Wafer, Chunky Bar, Marlboro Light.
9	Tango Wafer, Taro Net, Djava Young Bakery, Ohayo Bakery, Qtela, Putri Chania Pukis, Koko Krunch.
10	Aqua, Lasegar, Bear Brand, Buavita, Fanta Strawbery.
11	Aqua, Peyek Anugrah, Ohayo Bakery, Teh Pucuk, Lays, Kist Mint, Aim Biskuit Crispy, Putri Chania Pukis, Tebs Tea, Sari Roti, MM Nutriboost.
12	Cheetos, Ohayo Bakery, Aqua, Teh Botol, Kopiko, Hexos.
13	Sampurna 16, Proman Energenesis, Bear Brand.

14	Teh Pucuk, Monde Butter Cookies, Inaco Jelly, Fruit Tea, Indomie Kari Ayam, Biskuat Coklat.
15	Chitato, Lotte Koala's March, Koala March, Harapan Baru, Kacang Garuda, Mini Corneto Vanila, Silver Queen Black, M-150, Restu Bunda, Peyek Dua Putri, Chips Ahoy, Super Duluxe.
16	Aqua, Ohayo Bakery, Hexos, Abc Sari Kacang, Shanty Keripik Balado.
17	Populaire Stro-Van, PP.Rainbow Power, Lactogen, Clas Mild 16.
18	Lasegar, Harapan Baru, Kacang Klt Garuda Bawang, Charm Bf Extra Maxi 8, Walls Feast Chocolate, Paddle Pop Fruity, Cheetos, Orange Water, Chunky Bar, Fruit Tea, Teh Botol.
19	Fruit Tea, Kopiko, Rexona Men Ico Cool, Lays.
20	Sampurna 16, M-150.
21	Populaire Stro-Van, Sari Roti, Lays, Lara Kue Bawang, Mytea, Sampurna 16, Yazumi Bakery, Pulpy Orange, Silver Queen Silver, Aqua.
22	Hexos, Antimo, Paddle Pop Shaky Shake, Aqua.
23	Sari Roti, Paddle Pop Shaky Shake, Mytea.
24	Aqua, Teh Pucuk, Lays, Kopiko, You C1000 Orange, Pulpy Orange, Cheetos, Monde Butter Cookies, Fitbar Coklat, Paseo Smart Refil, Tango Wafer.
25	Sampurna 16, Tango Wafer, Niss Lemonia, Good Time Precious, Bear Brand, Lays, Kacang Garuda.
26	Rexona Ro, Orange Water.
27	Lays, Paddle Pop Ocean, Tango Wafer, PP.Choco Magma, Boom Pot Mint, Populaire Cup Choc-Vanila, Frestea Melati, Sweety XI 8 Extra, Faajar Kerupuk Bawang, Lolipop Pelangi Kecil.
28	Antangan Permen, Restu Bunda, Fajar Kerupuk Bawang, Pulpy Orange, Aqua.
29	Paddle Pop Color Paper, Paddle Pop Roket Jely.
30	Aqua, Garuda KCG Atom Manis, Lara Kue Bawang, Nescafe Latte, Chitato, Rifa Snack Kue Arai Pinang, Pulpy Orange, Sum Ayu S Sirih.
31	J.A Styl Cream, Vaseline Hbl, Axe 150, Kinder, Mb Lotion, Milna Biskuit, Pepsodent 120, Indomie Kari Ayam, Indomie Goreng, Indofood Sambal, Teh Botol, Ohayo Bakery.
32	Aqua, Paddle Pop Fruity Zap, PP.Rainbow Power.
33	Fajar Kerupuk Bawang, Peyek Anugrah, Snack Ball Jagung, Putri Chania Pukis, Teh Botol, Restu Bunda, Chesse Ring.
34	Clas Mild 16, Dunhil Filter 16, Garnier Light Spf.

Tabel 4.1 Tabel Transaksi

4.2 Analisa Pola Frekuensi Tinggi

Tahap ini mencari kombinasi *item* yang memenuhi syarat minimum dari nilai *support* dalam *database*. Nilai *support* sebuah *item* diperoleh dengan rumus berikut :

$$Support (A) = \frac{\text{Jumlah transaksi mengandung } A}{\text{Total Transaksi}}$$

Nilai *support* dari 2 item diperoleh dengan menggunakan rumus :

$$Support (A,B) = P(A \cap B)$$

$$Support (A,B) = \frac{\sum \text{Transaksi mengandung } A \text{ dan } B}{\sum \text{Transaksi}}$$

Setiap transaksi mempertasikan himpunan item yang berada dalam himpunan item yang dijual. Aturan yang kuat adalah aturan – aturan yang melebihi kriteria *support* dan *confidence* minimum. Misalnya seorang analisis menginginkan aturan yang memiliki *suppoort* lebih dari 10% dan *confidence* lebih dari 10%. Sebuah *itemset* adalah himpunan *item – item* yang ada dan *k-itemset* adalah *itemset* yang berisi *k-item* dan {A, B, C} merupakan 3 *itemset*.

Frequent itemset adalah *itemset* yang terjadi paling sedikit pada jumlah tertentu, ditentukan dengan banyak transaksi yaitu 34 dibagi dengan banyak item barang pada transaksi , sehingga dapat dicari dengan rumus :

$$Support (A) = \frac{\text{Jumlah transaksi mengandung } A}{\text{Total Transaksi}}$$

Untuk $k = 2$ (2 item), dari setiap item yang melewati *support* 30% akan dicari *frequent itemset* masing – masing item Nilai *support* dan 2 *item* diperoleh dengan menggunakan rumus :

$$Support (A,B) = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Trasaksi}}$$

Dari data yang sudah di masukan dan nilai dari minimum *support* yang telah ditentukan yaitu 10%, dan minimum *confidence* 10%, maka item yang mempunyai nilai minimum *support* diatas dari 10% dan nilai minimum *confidence* 10%, maka item tersebut yang akan di ambil untuk menentukan nilai *frequent itemset*.

4.3 Pembentukan Aturan Asosiasi

Setelah semua pada *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 dari rumus berikut:

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

$$Confidence = P(5|49) = \frac{\sum(\text{Aqua,Lays})}{\sum \text{Aqua}} = \frac{0,12}{0,41} = 0,29 = 29\%$$

$$Confidence = P(49|5) = \frac{\sum(\text{Lays,Aqua})}{\sum \text{Lays}} = \frac{0,12}{0,18} = 0,67 = 67\%$$

$$Confidence = P(5|67) = \frac{\sum(\text{Aqua,Ohayo Bakery})}{\sum \text{Aqua}} = \frac{0,12}{0,41} = 0,29 = 29\%$$

$$Confidence = P(67|5) = \frac{\sum(\text{Ohayo Bakery,Aqua})}{\sum \text{Ohayo Bakery}} = \frac{0,12}{0,21} = 0,57 = 57\%$$

$$Confidence = P(5|85) = \frac{\sum(\text{Aqua,Pulpy Orange})}{\sum \text{Aqua}} = \frac{0,18}{0,41} = 0,44 = 44\%$$

$$Confidence = P(85|5) = \frac{\sum(\text{Pulpy Orange,Aqua})}{\sum \text{Pulpy Orange}} = \frac{0,18}{0,21} = 0,86 = 86\%$$

$$Confidence = P(5|94) = \frac{\sum(\text{Aqua,Sari Roti})}{\sum \text{Aqua}} = \frac{0,12}{0,41} = 0,29 = 29\%$$

$$Confidence = P(94|5) = \frac{\sum(\text{Sari roti,Aqua})}{\sum \text{Sari roti}} = \frac{0,12}{0,15} = 0,80 = 80\%$$

$$Confidence = P(67|86) = \frac{\sum(\text{Ohayo Bakery,Putri Chania Pukis})}{\sum \text{Ohayo Bakeri}} = \frac{0,12}{0,21} = 0,57 = 57\%$$

$$Confidence = P(86|67) = \frac{\sum(\text{Putri Chania Pukis,Ohayo Bakery})}{\sum \text{Putri Chania Pukis}} = \frac{0,12}{0,15} = 0,80 = 80\%$$

Hasil pencarian untuk *confidence* yang didapat dari perhitungan kombinasi 2 *itemset* yang terdapat pada Tabel 4.2 berikut ini :

Knowledge	Confidence	Persentase
- Jika konsumen membeli Aqua, maka membeli Lays	0,29	29%
- Jika konsumen membeli Lays, maka membeli Aqua	0,67	67%
- Jika konsumen membeli Aqua, maka membeli Ohayo Bakery	0,29	29%
- Jika konsumen membeli Ohayo Bakery, maka membeli Aqua	0,57	57%
- Jika konsumen membeli Aqua, maka membeli Pulpy Orange	0,44	44%
- Jika konsumen membeli Pulpy Orange, maka membeli Aqua	0,86	86%
- Jika konsumen membeli Aqua, maka membeli Sari Roti	0,29	29%
- Jika konsumen membeli Sari Roti, maka membeli Aqua	0,80	80%
- Jika konsumen membeli Ohayo Bakery, maka membeli Putri Chania Pukis	0,57	57%
- Jika konsumen membeli Ohayo Bakery, maka membeli Jika konsumen membeli Putri Chania Pukis, maka membeli Ohayo Bakery	0,80	80%

Tabel 4.2 Tabel Knowledge

5 Kesimpulan

Dari penulisan tugas akhir ini mulai dari tahapan analisa permasalahan yang ada, hingga pengujian aplikasi sistem maka dapat diambil beberapa kesimpulan yaitu:

1. *Data Mining* dapat di implementasikan dengan menggunakan *database* penjualan produk barang untuk dapat menemukan kecenderungan pola kombinasi *itemsets* sehingga dapat dijadikan sebagai informasi untuk mengetahui prilaku konsumen dalam membeli produk barang secara bersamaan.
2. Aplikasi ini juga sebagai *alternative* alat bantu keputusan dalam menentukan penempatan barang diarea yang saling berdekatan sesuai prilaku konsumen dalam membeli barang secara bersamaan, membantu untuk mengetahui produk barang yang jarang dibeli konsumen dan sebagai alat *alternative*

dalam meningkatkan strategi pemasaran dengan cara membuat diskon barang tertentu yang jarang dibeli untuk menarik minat beli konsumen.

3. Dengan menerapkan aplikasi data mining menggunakan metode *association rules* ini dapat mengetahui bahwa produk yang paling tinggi lah persentase terjualnya maka produk tersebutlah yang akan ditingkatkan agar konsumen terpenuhi keinginannya.
4. Dengan menerapkan aplikasi data mining menggunakan metode *association rules* ini dapat mempermudah karyawan dalam memberikan laporan transaksi dengan cepat dan tepat terhadap pemimpinnya agar pimpinan dapat memberikan keputusan cepat dan tepat

Daftar Pustaka

- [1]. **Hermawati, Fajar Astuti.** "Data Mining. 2013." *ANDI: Yogyakarta*.
- [2]. **RIANGGA, JAYAPANA DUTA.** "Analisis Pola Pembelian Konsumen Dengan Algoritma Apriori Pada Data Transaksi Penjualan Obat Di Apotek Rahayu Jepara." *Skripsi, Fakultas Ilmu Komputer* (2015).
- [3]. **Ridwan, Mujib, Hadi Suyono, and M. Sarosa.** "Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier." *Jurnal EECCIS* : 59-64.
- [4]. **Pramusinto, Wahyu, and William Budi Utama Gunawan**-Universitas Budi. "Implementasi Algoritma Apriori untuk Aplikasi Data Mining Informasi Manfaat Asuransi Jiwa Studi Kasus: Pada PT Azarel Jelita Sejahtera." *Jurnal TICom* 2.1 (2013).
- [5]. Hadi, A. F. (2017). ANALISIS DATA MINING UNTUK MENENTUKAN VARIABEL-VARIABEL YANG MEMPENGARUHI KELAYAKAN KREDIT KEPEMILIKAN RUMAH MENGGUNAKAN TEKNIK KLASIFIKASI. *Komputer Teknologi Informasi*, 4(1).