

IMPLEMENTASI ALGORITMA APRIORI UNTUK MENENTUKAN POLA PEMBELIAN PADA CUSTOMER (STUDI KASUS : TOKO BAKOEL SEMBAKO)

Alfie Nur Rahmi¹⁾, Yosaphat Ananda Mikola²⁾

¹⁾ Sistem Informasi Universitas AMIKOM Yogyakarta

²⁾ Informatika Universitas AMIKOM Yogyakarta

email : alfienurrahmi@amikom.ac.id¹⁾, yosa.ananda14@gmail.com²⁾

Abstraksi

Bisnis ritel merupakan usaha yang sedang berkembang saat ini, semakin berkembangnya suatu usaha, maka transaksi yang terjadi sehari-hari pun semakin meningkat. Namun terkadang data transaksi tersebut belum dimanfaatkan secara maksimal oleh pemilik toko. Padahal, dengan menggunakan teknik data mining, kumpulan data tersebut dapat menghasilkan informasi baru, salah satunya adalah pola pembelian konsumen. Toko Bakoel Sembako merupakan salah satu toko ritel yang selama ini belum melakukan pemanfaatan data untuk penentuan strategi bisnis toko. Dengan adanya permasalahan tersebut, maka perlu adanya pemanfaatan data transaksi untuk diolah dengan algoritma apriori agar dapat memberikan pengetahuan baru yang dapat dimanfaatkan oleh pemilik toko. Berdasarkan hasil penelitian yang telah dilakukan, diperoleh hasil dari 30 data transaksi yang diolah, yang memenuhi minimal *support* 3 dan minimal *confidence* 70% diperoleh 1 aturan yang terbentuk yaitu jika pelanggan membeli produk Seedap mie goreng maka akan membeli telur 250gr.

Kata Kunci :

Data Mining, Algoritma Apriori, Algoritma

Abstract

The retail business is currently developing, the more a business develops, the transactions that occur every day are increasing. However, the transaction data has not been utilized optimally by the owner. In fact, by using data mining techniques, the data set can produce new information, one of which is consumer buying patterns. The Bakoel Sembako store is one of the retail stores that has not used data to determine the store's business strategy. With these problems, it is necessary to use transaction data to be processed with apriori algorithms in order to provide new knowledge that can be utilized by owners. Based on the results of research that has been carried out, the results obtained from 30 processed transaction data, which meet a minimum of support 3 and a minimum of 70% confidence, one rule is formed, namely if a customer buys Seedap mie goreng product, they will buy 250gr eggs.

Keywords :

Data Mining, Apriori Algorithm, Algorithm

1. Pendahuluan

Bisnis dan teknologi merupakan suatu hal yang tak terpisahkan, terutama saat pandemi seperti ini. semua sektor bisnis harus beralih menggunakan teknologi, salah satunya ada berjualan secara online. Bisnis ritel merupakan usaha yang sedang berkembang saat ini, semakin berkembangnya suatu usaha, maka transaksi yang terjadi sehari-hari pun semakin meningkat.

Setiap usaha ritel pasti memiliki catatan transaksi kegiatan operasinya, baik dicatat secara manual di buku atau yang sudah menggunakan komputer. Namun, nyatanya, data transaksi tersebut jarang dimanfaatkan oleh pemilik toko, padahal catatan transaksi tersebut dapat diolah kembali hingga menghasilkan informasi yang bermanfaat bagi pemilik toko. Salah satunya adalah dengan

pemanfaatan data mining. Data mining merupakan suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database [1].

Toko Bakoel Sembako merupakan salah satu contoh toko yang belum pernah memanfaatkan data transaksi yang mereka miliki. Mereka bahkan tidak tahu bahwa data transaksi yang mereka miliki dapat diolah menjadi sebuah informasi yang penting bagi strategi bisnis mereka. Berdasarkan hal tersebut, maka penulis akan melakukan penelitian penerapan algoritma apriori untuk menghasilkan pola pembelian customer pada Toko Bakoel Sembako dan diharapkan dari informasi tersebut dapat dimanfaatkan lebih lanjut oleh pemilik toko.

Penelitian terkait dengan Penerapan algoritma apriori telah banyak dilakukan, diantaranya adalah penelitian yang dilakukan oleh Novitasari yang menghasilkan

sebuah kesimpulan yaitu dengan adanya implementasi algoritma apriori sangat membantu Om Jeans dalam melihat pola penjualan yang banyak terjual, sehingga Om Jeans dapat menentukan dan memaksimalkan penyediaan stok barang sesuai permintaan konsumen [2]. Selain itu, penelitian yang dilakukan oleh Kurniawati, dkk menghasilkan kesimpulan bahwa penerapan data mining dengan menggunakan teknik algoritma apriori cukup efektif membantu dalam proses penentuan persediaan part pada perusahaan [3]. Selanjutnya, penelitian yang dilakukan oleh Aranda yakni mengimplementasikan metode *association rules* menggunakan algoritma apriori dari data mining, data presensi yang menumpuk tersebut dapat diasosiasikan dengan data lain untuk mendapatkan informasi dan pengetahuan tentang korelasi data presensi dengan data lain khususnya nilai yang berguna untuk optimalisasi jadwal kuliah. Optimalisasi yang diharapkan dari penelitian ini adalah meningkatkan kinerja atau prestasi mahasiswa dari pemilihan terbaik dari kombinasi jadwal kuliah, mata kuliah, dan dosen pengampu atau pun penemuan lainnya yang didapatkan dari hasil proses data mining [4].

Definisi 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. Pengelompokan data mining dijadikan sebagai teknik dari data mining berdasarkan tugas yang bisa dilakukan [5], yaitu:

1. Deskripsi

Para penulis biasanya mencoba menemukan cara untuk mendeskripsikan pola dan trend yang tersembunyi dalam data.

2. Klasifikasi

Dalam klasifikasi variabel, tujuan bersifat kategorik

3. Estimasi

Estimasi mirip dengan klasifikasi, kecuali variabel tujuan yang lebih ke arah numerik daripada kategori.

4. Prediksi

Prediksi memiliki kemiripan dengan estimasi dan klasifikasi. Hanya saja, prediksi hasilnya menunjukkan sesuatu yang belum terjadi (mungkin terjadi dimasa depan).

5. Clustering

Clustering lebih kearah pengelompokan record, pengamatan, atau kasus dalam kelas yang memiliki kemiripan.

6. Asosiasi

Mengidentifikasi hubungan antara berbagai peristiwa yang terjadi pada satu waktu.

Analisis asosiasi atau *association rule mining* adalah teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item. Analisis asosiasi dikenal juga sebagai salah satu teknik data mining

yang menjadi dasar dari berbagai teknik data mining lainnya. Secara khusus, salah satu tahap analisis asosiasi yang menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien adalah analisis pola frekuensi tinggi (*frequent pattern mining*). Penting tidaknya suatu aturan asosiasi dapat diketahui dengan dua parameter, yaitu *support* dan *confidence*. *Support* (nilai penunjang) adalah persentase kombinasi item tersebut dalam *database*, sedangkan *confidence* (nilai kepastian) adalah kuatnya hubungan antar-item dalam aturan asosiasi [6].

Apriori secara umum adalah mencari *frequent itemset* dari *database* transaksi penjualan melalui beberapa tahap iterasi. Langkah pertama algoritma ini adalah menghitung kejadian item untuk menentukan *large 1-itemset* atau *frequent itemset*. Langkah berikutnya untuk *large itemset* tersebut dilakukan penggabungan dari *item* yang telah lolos seleksi kemudian *itemset* yang tidak lolos maka akan dihapus. Langkah berikutnya untuk *item* yang telah lolos diiterasi tersebut maka akan digunakan untuk proses selanjutnya, dan yang tidak lolos akan dihapus. Demikian langkah demi langkah pencarian kandidat melalui iterasi sehingga memperoleh titik akhir yang tidak dapat dilakukan iterasi lagi [7].

Lift Ratio adalah suatu ukuran (parameter) untuk mengetahui kekuatan aturan asosiasi (*association rule*) yang telah terbentuk dari nilai *support* dan *confidence*. Nilai *lift ratio* biasanya digunakan sebagai penentu apakah aturan asosiasi valid atau tidak valid.

Cara kerja metode ini adalah membagi *confidence* dengan *expected confidence*. *Confidence* dapat dihitung dengan rumus :

$$Confidence P(B|A) = \frac{\text{Jumlah Transaksi mengandung A dan B}}{\text{Jumlah Transaksi mengandung A}} \quad (1)$$

Antecedent merupakan sebab yang menjadikan *item consequent*. Sedangkan *Consequent* adalah sebuah akibat atau juga *item* yang akan dibeli setelah membeli *Antecedent*. Nilai dari *expected confidence* dapat dihitung dengan rumus :

$$Expected\ confidence = \frac{\text{Jumlah Transaksi mengandung B}}{\text{Jumlah Transaksi}} \quad (2)$$

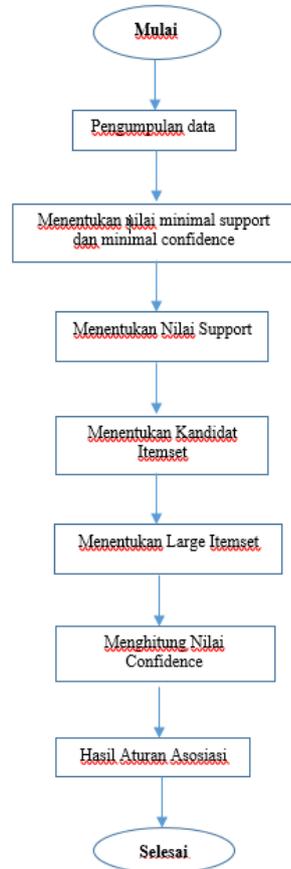
Lift ratio dapat dihitung dengan cara membandingkan antara *confidence* untuk suatu aturan dibagi dengan *expected confidence*. Berikut rumus dari *Lift ratio*:

$$Lift\ Ratio = \frac{Confidence}{Expected\ confidence} \quad (3)$$

Nilai *Lift ratio* lebih besar dari 1 menunjukkan adanya manfaat dari aturan tersebut. Lebih tinggi nilai *Lift ratio*, lebih besar kekuatan asosiasinya [8].

2. Metode Penelitian

Berikut alur pada penelitian ini :



Gambar 1 Alur penelitian

Langkah awal penelitian dimulai dari pengumpulan data, lalu dari data yang dikumpulkan selanjutnya penulis menentukan nilai minimal *support* dan minimal *confidence* yang akan digunakan untuk mencari aturan asosiasinya. Selanjutnya menentukan nilai *support* dan kandidat *itemset*. Dari kandidat *itemset* yang telah ada, dipilih yang memenuhi minimal *support* untuk menjadi *large itemset*. Setelah itu menghitung nilai *confidence* hingga akhirnya ditemukan hasil aturan asosiasi nya.

3. Hasil dan Pembahasan

Dalam perhitungan apriori terlebih dahulu mengumpulkan data transaksi penjualan di Toko Bakoel Sembako sebanyak 30 transaksi secara acak pada 1 Desember sampai 6 Desember 2020. Data yang diambil minimal 2 penjualan produk. Data transaksi yang akan dihitung ditunjukkan pada tabel 1 :

Tabel 1 Data Transaksi

Id	Tanggal	Produk
1	01/12/2020	telur 250gr, sedaap mie goreng
2	01/12/2020	telur 250gr, minyak sawit 500gr
3	01/12/2020	sedaap kecap manis 550ml, gula pasir 500g
4	01/12/2020	mie burung dara kuning(136gr), sosis sonice sapi/ayam
5	01/12/2020	telur 250gr, minyak sawit 500gr
6	01/12/2020	royco sapi (9gr), mie burung dara kuning(136gr)
7	01/12/2020	teh tjatoet, teh dandang biru 40g
8	02/12/2020	bihun padamu 350g, minyak sawit 500gr
9	02/12/2020	kapal api 65g, gula pasir 500g
10	02/12/2020	indomie goreng special, sosis sonice sapi/ayam
11	02/12/2020	indomie goreng hype abis, migor ayam geprek, indomie goreng special
12	02/12/2020	minyak sawit 500gr, bihun padamu 350g
13	02/12/2020	desaku ketumbar bubuk 15g, desaku kunyit bubuk 10g
14	03/12/2020	gula pasir 500g, sosis sonice sapi/ayam
15	03/12/2020	indomie goreng special, indomie ayam special
16	03/12/2020	desaku kunyit bubuk 10g, desaku ketumbar bubuk 15g
17	03/12/2020	indomie ayam special, indomie goreng special
18	03/12/2020	kapal api 65g, sedaap kecap manis 225ml
19	03/12/2020	kapal api 65g, gula pasir 500g
20	04/12/2020	sedaap mie goreng, telur 250gr
21	04/12/2020	minyak sawit 500gr, royco sapi (9gr)
22	04/12/2020	ladaku (4gr), royco ayam 9g
23	05/12/2020	teh tjatoet, gula pasir 500g

Tabel 1 Lanjutan

Id	Tanggal	Produk
24	05/12/2020	teh dandang biru 40g, gula pasir 500g
25	05/12/2020	minyak sawit 500gr, gula pasir 500g
26	05/12/2020	ladaku (4gr), royco ayam 9g
27	05/12/2020	desaku ketumbar bubuk 15g, ladaku (4gr)
28	06/12/2020	sedaap mie goreng, telur 250gr
29	06/12/2020	sosis sonice sapi/ayam, frestea jasmine 296 ml
30	06/12/2020	sedaap kecap manis 225ml, minyak sawit 500gr

Selanjutnya menentukan nilai minimal *support* dan minimal *confidence* untuk mencari aturan asosiasinya. Nilai minimal *support* yang ditentukan adalah 3 dan minimal *confidence* adalah 70%. Pertama-tama kita akan mencari *itemset* 1 (C1) dengan menghitung jumlah frekuensi tiap produk beserta nilai *support* nya. Cara menghitung nilai *support* dalam persen dan kandidat *Itemset* 1 di tunjukan pada tabel 2 :

$$Support(A) = \frac{Jumlah\ Transaksi\ mengandung\ A}{Jumlah\ Transaksi} \times 100\% \quad (4)$$

$$Minimal\ Support = \frac{3}{30} \times 100\% = 10$$

$$Support(Telur\ 250gr) = \frac{5}{30} \times 100\% = 16.67$$

Tabel 2 Kandidat *Itemset* 1

No	Item	Jml	Support	Ket.
1	telur 250gr	5	16.67	lolos
2	Sedaap mie goreng	3	10.00	lolos
3	minyak sawit 500gr	7	23.33	lolos
4	sedaap kecap manis 550ml	1	3.33	tidak lolos
5	gula pasir 500g	7	23.33	lolos
6	mie burung dara kuning(136gr)	2	6.67	tidak lolos

Tabel 2 Lanjutan

No	Item	Jml	Support	Ket.
7	sosis sonice sapi/ayam	4	13.33	lolos
8	royco sapi (9gr)	2	6.67	tidak lolos
9	teh tjatoet	2	6.67	tidak lolos
10	teh dandang biru 40g	2	6.67	tidak lolos
11	bihun padamu 350g	2	6.67	tidak lolos
12	kapal api 65g	3	10.00	lolos
13	indomie goreng special	4	13.33	lolos
14	indomie goreng hype abis migor ayam geprek	1	3.33	tidak lolos
15	desaku ketumbar bubuk 15g	3	10.00	lolos
16	desaku kunyit bubuk 10g	2	6.67	tidak lolos
17	indomie ayam special	2	6.67	tidak lolos
18	sedaap kecap manis 225ml	2	6.67	tidak lolos
19	ladaku (4gr)	3	10	lolos
20	royco ayam 9g	2	6.67	tidak lolos
21	frestea jasmine 296 ml	1	3.33	tidak lolos

Langkah selanjutnya adalah menyeleksi kandidat *itemset* yang tidak memenuhi minimal *support* yang ditentukan, yaitu 3. Tiap kandidat *itemset* 1 yang lolos akan menjadi *large itemset* (L1) yang di tunjukan pada tabel 3.

Tabel 3 Large Itemset 1 (L1)

No	Item	Jml	Support
1	telur 250gr	5	16.67
2	sedaap mie goreng	3	10.00
3	minyak sawit 500gr	7	23.33
4	gula pasir 500g	7	23.33
5	sisis sonice sapi/ayam	4	13.33
6	kapal api 65g	3	10.00
7	indomie goreng special	4	13.33
8	desaku ketumbar bubuk 15g	3	10.00
9	ladaku (4gr)	3	10.00

Tabel 4 Lanjutan

No	Item1	Item2	Jml	Support	Ket.
8	telur 250gr	ladaku (4gr)	0	0.00	tidak lolos
9	sedaap mie goreng	minyak sawit 500gr	0	0.00	tidak lolos
10	sedaap mie goreng	gula pasir 500g	0	0.00	tidak lolos
11	sedaap mie goreng	sisis sonice sapi/ayam	0	0.00	tidak lolos
12	sedaap mie goreng	kapal api 65g	0	0.00	tidak lolos
13	sedaap mie goreng	indomie goreng special	0	0.00	tidak lolos
14	sedaap mie goreng	desaku ketumbar bubuk 15g	0	0.00	tidak lolos
15	sedaap mie goreng	ladaku (4gr)	0	0.00	tidak lolos
16	minyak sawit 500gr	gula pasir 500g	1	3.33	tidak lolos
17	minyak sawit 500gr	sisis sonice sapi/ayam	0	0.00	tidak lolos
18	minyak sawit 500gr	kapal api 65g	0	0.00	tidak lolos
19	minyak sawit 500gr	indomie goreng special	0	0.00	tidak lolos
20	minyak sawit 500gr	desaku ketumbar bubuk 15g	0	0.00	tidak lolos
21	minyak sawit 500gr	ladaku (4gr)	0	0.00	tidak lolos
22	gula pasir 500g	sisis sonice sapi/ayam	1	3.33	tidak lolos
23	gula pasir 500g	kapal api 65g	2	6.67	tidak lolos
24	gula pasir 500g	indomie goreng special	0	0.00	tidak lolos
25	gula pasir 500g	desaku ketumbar bubuk 15g	0	0.00	tidak lolos
26	gula pasir 500g	ladaku (4gr)	0	0.00	tidak lolos
27	sisis sonice sapi/ayam	kapal api 65g	0	0.00	tidak lolos
28	sisis sonice sapi/ayam	indomie goreng special	1	3.33	tidak lolos
29	sisis sonice sapi/ayam	desaku ketumbar bubuk 15g	0	0.00	tidak lolos

Setelah mendapatkan *large itemset*, maka iterasi kedua dilakukan dengan menghitung 2 kombinasi *item* untuk menghasilkan kandidat *itemset 2 (C2)*. Perhitungan *support* tiap *itemset* dilakukan dengan persamaan yang sama dalam mencari *itemset 1* sebelumnya, adapun cara melakukan perhitungan *itemset 2* sebagai berikut:

$$Support(A, B) = \frac{Jumlah\ Transaksi\ mengandung\ A,B}{Jumlah\ Transaksi} \times 100\%$$

$$Support(Telur\ 250gr, Sedaap\ Mie\ Goreng) = \frac{3}{30} \times 100\% = 10$$

Tabel 4 Kandidat Itemset 2 (C2)

No	Item1	Item2	Jml	Support	Ket.
1	telur 250gr	sedaap mie goreng	3	10.00	lolos
2	telur 250gr	minyak sawit 500gr	2	6.67	tidak lolos
3	telur 250gr	gula pasir 500g	0	0.00	tidak lolos
4	telur 250gr	sisis sonice sapi/ayam	0	0.00	tidak lolos
5	telur 250gr	kapal api 65g	0	0.00	tidak lolos
6	telur 250gr	indomie goreng special	0	0.00	tidak lolos
7	telur 250gr	desaku ketumbar bubuk 15g	0	0.00	tidak lolos

Tabel 4 Lanjutan

No	Item1	Item2	Jml	Support	Ket.
30	sosis sonice sapi/ayam	ladaku (4gr)	0	0.00	tidak lolos
31	kapal api 65g	indomie goreng special desaku	0	0.00	tidak lolos
32	kapal api 65g	ketumbar bubuk 15g	0	0.00	tidak lolos
33	kapal api 65g	ladaku (4gr) desaku	0	0.00	tidak lolos
34	indomie goreng special	ketumbar bubuk 15g	0	0.00	tidak lolos
35	indomie goreng special	ladaku (4gr)	0	0.00	tidak lolos
36	indomie goreng special	ladaku (4gr)	1	3.33	tidak lolos

Selanjutnya kembali memangkas *itemset* yang memiliki nilai minimal *support* 3 yang di tunjukan pada tabel 5.

Tabel 5 Large Itemset 2 (L2)

No	Item 1	Item 2	Jml	Support
1	telur 250gr	sedaap mie goreng	3	10.00

Iterasi ke 3 untuk mendapatkan kandidat *itemset* ke 3 (C3) tidak dapat dilakukan karena tidak ada yang memenuhi minimal *support*, maka iterasi akan berhenti. Setelah menentukan *large itemset*, kemudian dilakukan perhitungan bagian *confidence* dalam bentuk persen dan hasilnya ditunjukkan pada tabel 6.

Tabel 6 Hasil Perhitungan Nilai Confidence

No	A => B	Support A U B	Support A	Confidence	Ket.
1	sedaap mie goreng => telur 250gr	10.00	10.00	100.00	lolos
2	telur 250gr => sedaap mie goreng	10.00	16.67	60.00	tidak lolos

Setelah nilai *confidence* diperoleh, langkah selanjutnya mengeleminasi *itemset* yang tidak memenuhi minimal *confidence* yaitu 70%. Hasil ini

merupakan hasil akhir dari aturan asosiasi yang di tunjukan pada tabel 7

Item1	Item2	Jml	Support	Ket.
Tabel 7 Hasil Aturan Asosiasi				
no	x => y	confidence	nilai uji lift	korelasi rule
1	sedaap mie goreng => telur 250gr	100.00	6.00	korelasi positif

Dari 30 transaksi yang diuji menghasilkan 1 aturan asosiasi yang memenuhi minimal *support* 3 dan minimal *confidence* 70%. Aturan yang terbentuk adalah jika pelanggan membeli produk Sedaap mie goreng maka akan membeli telur 250 gr.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat diambil kesimpulan bahwa dari 30 data transaksi yang diolah, yang memenuhi minimal *support* 3 dan minimal *confidence* 70% diperoleh 1 aturan yang terbentuk yaitu jika pelanggan membeli produk Sedaap mie goreng maka akan membeli telur 250 gr. Setelah diketahui hasilnya, informasi tersebut dapat digunakan oleh pemilik toko untuk menyusun strategi bisnisnya, bisa untuk strategi penentuan stok artinya jika menambah stok Sedaap Mie Goreng, maka stok telur 250gr juga ditambah, atau bisa juga digunakan untuk strategi tata letak penempatan barang, missal Sedaap mie goreng diletakkan berdampingan dengan telur 250gr atau untuk keperluan promo barang, misalkan promo bundling Sedaap mie goreng dengan telur 250gr.

Daftar Pustaka

- [1] E. Turban, Decision Support System and Intellegent Systems, Yogyakarta: Andi,2005.
- [2] H. Novitasari, Implementasi Algoritma Apriori dalam Menentukan Pola Penjualan Produk pada Om Jean, Yogyakarta: Universitas Amikom,2019.
- [3] L. Kurniawati, A. E. Kusuma, B. Dewansyah, Implementasi Algoritma Apriori untuk Menentukan Persediaan Spare Part Compressor, Journal of Computer Engineering System and Science (CESS) Vol. 4. p-ISSN :2502-7131,2019.
- [4] J. Aranda, Implementasi Data Mining Algoritma Apriori Untuk Optimalisasi Jadwal Kuliah yang Mempengaruhi Nilai Akhir Mahasiswa Per Mata Kuliah dengan Studi kasus di Baak Stmik Amikom Yogyakarta, Yogyakarta: Universitas Amikom,2016.
- [5] Kusri, Luthfi, Algoritma Data Mining, Yogyakarta: Andi Offset,2009.
- [6] I. Ukiarwan, Penentuan Aturan Asosiasi Pada Penjualan Produk Sepatu Running Menggunakan Algoritma Apriori (Studi Kasus : Jogja Sepatu), Yogyakarta: Universitas Sanata Dharma,2017.
- [7] A. D. Setyawan, Aplikasi Data Mining Menggunakan Algoritma Apriori untuk Analisis Pola Penjualan Di Xyz Helm, Yogyakarta: Universitas Sanata Dharma,2016.
- [8] B. Santosa, Data Mining : Teknik Pemanfaatan Data untuk Keperluan Bisnis, Yogyakarta: Graha Ilmu,2007.