

Penerapan Association Rule Menggunakan Frequent Pattern Growth Untuk Rekomendasi Produk Jersey Sepakbola

Anwar Musaddad¹, Odi Nurdiawan^{2,*}, Gifthera Dwilestari³

¹ Program Studi Teknik Informatika, STMIK IKMI, Cirebon, Indonesia

² Program Studi Manajemen Informatika, STMIK IKMI, Cirebon, Indonesia

³ Program Studi Sistem Informasi, STMIK IKMI, Cirebon, Indonesia

Email: ¹anwarmusaddad007@gmail.com, ^{2,*}odinurdiawan2020@gmail.com, ³ggdwilestari@gmail.com

Email Penulis Korespondensi: odinurdiawan2020@gmail.com

Submitted: 09/03/2022; Accepted: 25/05/2022; Published: 31/05/2022

Abstrak—Fenomena awal tahunan hal yang ditunggu oleh beberapa fans sepak bola yaitu terbitnya jersey terbaru dari tim kesayangannya. Ketika jersey baru sudah launching fans sepak bola berbondong-bondong untuk membeli jersey tersebut akan tetapi ada beberapa toko yang tersedia jersey terbarunya. Hal tersebut dialami oleh toko eighteen sport dalam memenuhi keinginan fans terdapat kendala re-stok ulang jersey yang paling diminati. Sehingga banyak barang yang belum laku terjual. Fokus penelitian ini terletak pada pengelolaan data penjualan jersey pada bulan juni, juli dan agustus, serta minat yang tinggi akan permintaan jersey club. Permintaan jersey yang tinggi di pengaruhi oleh capaian dari klubnya itu sendiri. Penelitian ini menggunakan metode algoritma FP Growth dengan tujuan mendapatkan pola rekomendasi dari keinginan fans bola. Hasil penelitian yang telah dilakukan penulis dapat menarik beberapa kesimpulan yaitu Berdasarkan hasil pengelolaan support didapatkan bahwa konsumen dengan membeli 1 item jersey akan membeli kembali 1 item jersey yang berbeda sebanyak 15 Pola. Konsumen dengan membeli 2 item jersey akan membeli kembali 1 item jersey yang berbeda sebanyak 46 Pola. Konsumen dengan membeli 3 item jersey akan membeli kembali 1 item jersey yang berbeda sebanyak 37 Pola. Konsumen dengan membeli 4 item jersey akan membeli kembali 1 item jersey yang berbeda sebanyak 10 Pola. Sehingga data pol tersebut menjadi rekomendasi owner untuk melakukan pembelian ulang.

Kata Kunci: Jersey Bola; Fans Bola; Algoritma FP Growth

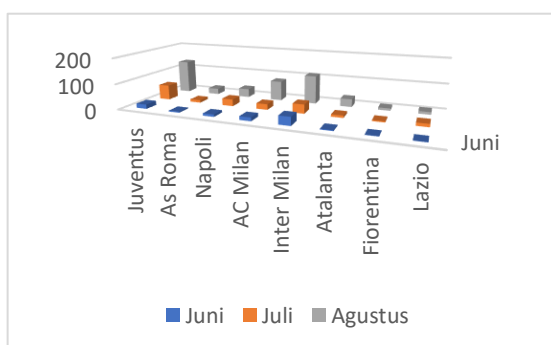
Abstract—The phenomenon of the beginning of the year, what some football fans have been waiting for, is the publication of the latest jersey from their favorite team. When the new jersey was launched, football fans flocked to buy the jersey, but there were several shops available for the new jersey. This was experienced by the Eighteen Sport shop, in fulfilling the wishes of fans, there were obstacles to re-stock the jerseys that were most in demand. So many items that have not been sold. The focus of this research lies in managing jersey sales data in June, July and August, as well as high interest in the demand for club jerseys. The high demand for jerseys is influenced by the achievements of the club itself. This study uses the FP Growth algorithm with the aim of getting a recommendation pattern from the wishes of football fans. Based on the results of the support management, it was found that consumers by buying 1 jersey item will buy back 1 different jersey item as many as 15 patterns. Consumers by buying 2 jersey items will repurchase 1 different jersey item as many as 46 patterns. Consumers by buying 3 jersey items will repurchase 1 different jersey item as many as 37 patterns. Consumers by buying 4 jersey items will repurchase 1 different jersey item for 10 patterns. So that the pol data becomes the owner's recommendation to make a repeat purchase.

Keywords: Soccer Jersey; Football Fans; FP Growth Algorithm

1. PENDAHULUAN

Seiring perkembangan zaman dibidang teknologi, menuntut perusahaan atau pelaku usaha untuk menggunakan teknologi sebagai media untuk tetap bertahan dan memenangkan persaingan bisnis yang ketat. Perusahaan harus memiliki kemampuan untuk beradaptasi terhadap perubahan yang terjadi sehingga mampu bersaing dengan para kompetitor pasar. Saat ini internet menjadi salah satu media komunikasi terbesar, dengan internet semua orang didunia bisa mendapatkan dan menyampaikan informasi yang dibutuhkan kapan saja dan dimana saja. Kini dengan hadirnya internet, manusia dapat melakukan bisnis lebih mudah.

Sepakbola adalah sebuah permainan olahraga yang dimainkan dua tim dan masing-masing memiliki sebelas orang pemain dan saling bertanding untuk memasukan bola ke dalam gawang. Kesederhanaan bermain dan peraturan yang mudah dipahami menjadi salah satu faktor olahraga ini populer di seluruh lapisan masyarakat.



Gambar 1. Grafik Penjualan

Fenomena awal tahunan hal yang ditunggu oleh beberapa fans sepak bola yaitu terbitnya jersey terbaru dari tim kesayangannya. Ketika jersey baru sudah launching fans sepak bola berbondong-bondong untuk membeli jersey tersebut akan tetapi ada beberapa toko yang tersedia jersey terbarunya. Hal tersebut dialami oleh toko Anugrah Sport dalam memenuhi keinginan fans terdapat kendala re-stok ulang jersey yang paling diminati. Sehingga banyak barang yang belum laku terjual.

Fokus penelitian ini terletak pada pengelolaan data penjualan jersey pada bulan juni juli dan agustus, serta minat yang tinggi akan permintaan jersey club. Permintaan jersey yang tinggi di pengaruhi oleh capaian dari klubnya itu sendiri. Penelitian ini menggunakan metode algoritma *Frequent Pattern Growth* dengan tujuan mendapatkan pola rekomendasi dari keinginan fans bola.

Frequent Pattern Growth merupakan salah satu alternative algoritma yang bisa digunakan dalam penentuan himpunan data yang paling sering muncul (*frequent itemset*) pada sebuah kumpulan data (Riarianti, 2014). Untuk mencari *frequent itemset* dengan menggunakan *frequent Pattern growth* adalah dengan cara menggunakan FP-tree yang merupakan perluasan dari pohon prefix. Langkah pertama yang dilakukan yaitu pembentukan *frequent Pattern tree*. Setelah pembentukan *frequent Pattern tree* dapat diterapkan algoritma *frequent Pattern growth* untuk langsung mengekstrak frequent itemset dari *frequent Pattern tree* yang telah ada dengan menggunakan prinsip *divide and conquer*.

2. METODOLOGI PENELITIAN

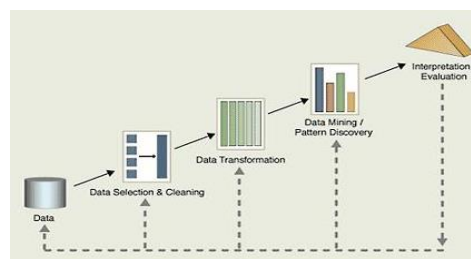
Pada penelitian ini peneliti menggunakan pendekatan kuantitatif. Dalam pengertiannya penelitian deskriptif adalah penelitian yang menggambarkan suatu keadaan atau kondisi yang terjadi disuatu daerah, yang pemecahan masalahnya secara tersistem berdasarkan data – data yang bersikap fakta yang ada. dan data kuantitatif adalah jenis data berupa angka yang dapat digunakan dalam proses operasi matematika, data ini berasal dari data fakta yang ada. Jenis data yang digunakan dalam penelitian ini bersifat rasio yang merupakan data yang berbentuk angka sebenarnya.

2.1 Teknik Pengumpulan Data

Adapun teknik pengumpulan data merupakan suatu cara yang dilakukan oleh peneliti untuk memperoleh data-data yang diperlukan. Dalam penelitian ini teknik yang digunakan adalah survey penelitian dengan meminta ijin pemilik toko, dan menggunakan sumber data Sekunder penjualan toko bulan Juni 2021 sampai dengan Agustus 2021 dan data Primer yang di dapatkan dengan wawancara

2.2 Tahapan Penelitian

Metode pengembangan data mining yang digunakan untuk menganalisis data dalam penerapan data mining ini menggunakan proses tahapan *knowledge discovery in databases* (KDD) yang terdiri dari Data, Data Cleaning, Data transformation, Data mining, *Pattern evolution*, *knowledge* :



Gambar 2. Tahapan Penelitian

Berikut merupakan hal – hal yang perlu di lakukan dalam penelitian berdasarkan tahapan *knowledge discovery in databases*;

1. Data

Data merupakan sekumpulan data operasional yang diperlu sebelum dilakukan sebelum tahap penggalian informasi dalam *Knowledge Discovery Database* (KDD) dimulai.

2. Data Cleaning

Proses *data cleaning* merupakan proses Pembersihan data yang bertujuan untuk menghilangkan data yang tidak memiliki nilai (null), data yang salah input, data yang tidak relevan, duplikat data dan data yang tidak konsisten karena keberadaanya bisa mengurangi mutu atau akurasi dari hasil data mining nantinya. Pembersihan data juga akan mempengaruhi performasi dari system data mining karena data yang akan ditangani akan berkurang jumlah dan kompleksitasnya.

3. Data transformation

Data transformation dilakukan dengan memberikan inialisasi terhadap data yang memiliki nilai nominal menjadi bernilai numerik.

4. Data Mining

Pada fase ini yang dilakukan adalah menerapkan algoritma atau metode pencarian pengetahuan. Ini adalah langkah penting di mana teknik kecerdasan diterapkan untuk mengekstrak pola informasi yang berpotensi berguna dari data yang dipilih.

Data Mining adalah proses mencari pola atau informasi menarik dalam data terpilih menggunakan teknik atau metode tertentu. Salah satu metode yang seringkali digunakan dalam Data Mining adalah metode Association Rule. Algoritma FP-Growth merupakan pengembangan dari algoritma Apriori. Algoritma Frequent Pattern Growth adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (frequent itemset) dalam sebuah kumpulan data [5].

Pada algoritma FP-Growth menggunakan konsep pembangunan tree, yang biasa disebut FP-Tree, dalam pencarian frequent itemsets bukan menggunakan generate candidate seperti yang dilakukan pada algoritma Apriori. Dengan menggunakan konsep tersebut, algoritma FP-Growth menjadi lebih cepat daripada algoritma Apriori. [2] Metode FP-Growth dibagi menjadi tiga tahapan utama, yaitu [4]:

- Tahap pembangkitan *conditional pattern base*,
- Tahap pembangkitan *conditional FP-Tree*, dan
- Tahap pencarian *frequent itemset*.

Association rule merupakan suatu proses pada data mining untuk menentukan semua aturan asosiatif yang memenuhi syarat minimum untuk support (minsup) dan confidence (minconf) pada sebuah database. Kedua syarat tersebut akan digunakan untuk interesting association rules dengan dibandingkan dengan batasan yang telah ditentukan, yaitu minsup dan minconf [3]. Association Rule Mining adalah suatu prosedur untuk mencari hubungan antar item dalam suatu dataset. Dimulai dengan mencari frequent itemset, yaitu kombinasi yang paling sering terjadi dalam suatu itemset dan harus memenuhi minsup [3]. Dalam tahap ini akan dilakukan pencarian kombinasi item yang memenuhi syarat minimum dari nilai support dalam database. Untuk mendapatkan nilai support dari suatu item A dapat diperoleh dengan rumus berikut: [1]

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

Kemudian, untuk mendapatkan nilai *support* dari dua item diperoleh dengan rumus berikut:

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

Setelah semua *frequent item* dan *large item set* didapatkan, dapat dicari syarat *minimum confidence (minconf)* dengan menggunakan rumus berikut :

$$\text{Confidence (A} \rightarrow \text{B)} = \frac{P(A|B) \text{ Jumlah Transaksi yang mengandung A dan B}}{\text{Jumlah Transaksi yang mengandung A}} \quad (3)$$

5. Evaluation

Pada tahap evaluasi akan diketahui apakah hasil daripada tahap data mining dapat menjawab tujuan yang telah ditetapkan. Untuk itu akan dilakukan profilisasi pada setiap cluster yang telah terbentuk, untuk diketahui karakteristik pada kelompok tersebut. Disamping itu untuk diketahui kesesuaian dengan jalur perminatan akan dilakukan analisis lebih lanjut untuk dihubungkan dengan atribut perminatan, Sehingga diharapkan mendapatkan informasi atau pola yang berguna sebagai acuan pematkhiran data.

6. Knowledge

Tahap terakhir dari proses data mining adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisis yang didapat

3. HASIL DAN PEMBAHASAN

3.1 Data

Penelitian ini menggunakan data Skunder artinya data yang akan digunakan yaitu data nasabah pada bulan Juni Juli dan Agustus tahun 2021 dengan data penjualan dilihat pada tabel berikut ini :

Tabel 1. Data Penjualan

No	Tanggal	Bulan	Tahun	Atalanta	Bologna	Cagliari	Empoli	...	Fiorentina
1	1	Juni	2021	0	3	2	0	...	3
2	2	Juni	2021	0	1	0	1	...	1
3	3	Juni	2021	1	2	0	0	...	2
4	4	Juni	2021	0	4	1	0	...	4
5	5	Juni	2021	0	1	0	0	...	1
6	6	Juni	2021	0	0	0	0	...	0
7	7	Juni	2021	0	0	0	2	...	0
8	8	Juni	2021	0	2	0	0	...	2

No	Tanggal	Bulan	Tahun	Atalanta	Bologna	Cagliari	Empoli	...	Fiorentina
9	9	Juni	2021	0	5	0	0	...	5
10	10	Juni	2021	2	0	0	0	...	0
11	11	Juni	2021	0	2	2	3	...	2
12	12	Juni	2021	1	2	0	0	...	2
13	13	Juni	2021	0	0	1	0	...	0
14	14	Juni	2021	0	0	0	0	...	0
15	15	Juni	2021	0	0	0	0	...	0
16	16	Juni	2021	0	0	1	0	...	0
17	17	Juni	2021	0	1	0	0	...	1
18	18	Juni	2021	1	0	0	0	...	1
19	19	Juni	2021	1	0	2	0	...	0
...
123	31	Agustus	2021	0	0	0	1	...	0

3.2 Data Preprocessing

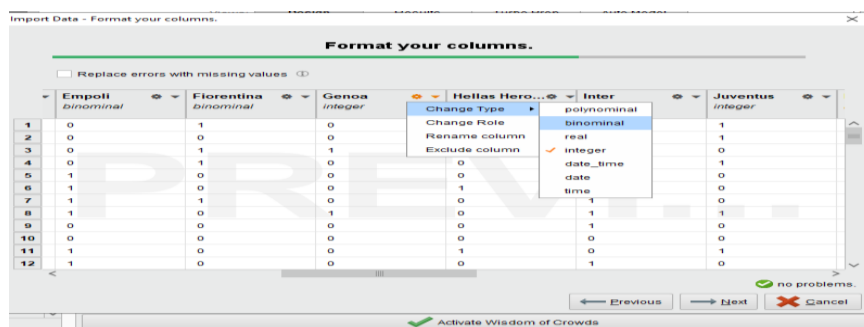
Pre-Processing merupakan tahap awal dataset akan dilakukan pembersihan data, data yang akan dibersihkan yakni data transaksi penjualan furnitur. Atribut yang dilakukan transformasi yakni atribut item, proses transformasi dilakukan dengan membuat atribut baru item menjadi atribut khusus nama-nama *item*/produk yang menjadi *record* data pada atribut sebelumnya, lalu memisahkan data yang lebih dari satu item dan melakukan pengisian data pada atribut – atribut baru dengan mengisi 1 jika melakukan transaksi dan 0 jika tidak melakukan transaksi.

Tabel 2 Hasil Preprocessing data

No	Tanggal	Bulan	Tahun	Atalanta	Bologna	Cagliari	Empoli	...	Fiorentina
1	1	Juni	2021	0	1	1	0	...	1
2	2	Juni	2021	0	1	0	1	...	1
3	3	Juni	2021	1	1	0	0	...	1
4	4	Juni	2021	0	1	1	0	...	1
5	5	Juni	2021	0	1	0	0	...	1
6	6	Juni	2021	0	0	0	0	...	0
7	7	Juni	2021	0	0	0	1	...	0
8	8	Juni	2021	0	1	0	0	...	1
9	9	Juni	2021	0	1	0	0	...	1
10	10	Juni	2021	1	0	0	0	...	0
11	11	Juni	2021	0	1	1	1	...	1
12	12	Juni	2021	1	1	0	0	...	1
13	13	Juni	2021	0	0	1	0	...	0
14	14	Juni	2021	0	0	0	0	...	0
15	15	Juni	2021	0	0	0	0	...	0
16	16	Juni	2021	0	0	1	0	...	0
17	17	Juni	2021	0	1	0	0	...	1
18	18	Juni	2021	1	0	0	0	...	1
19	19	Juni	2021	1	0	1	0	...	0
...
123	31	Agustus	2021	0	0	0	1	...	0

3.3 Data Transformation

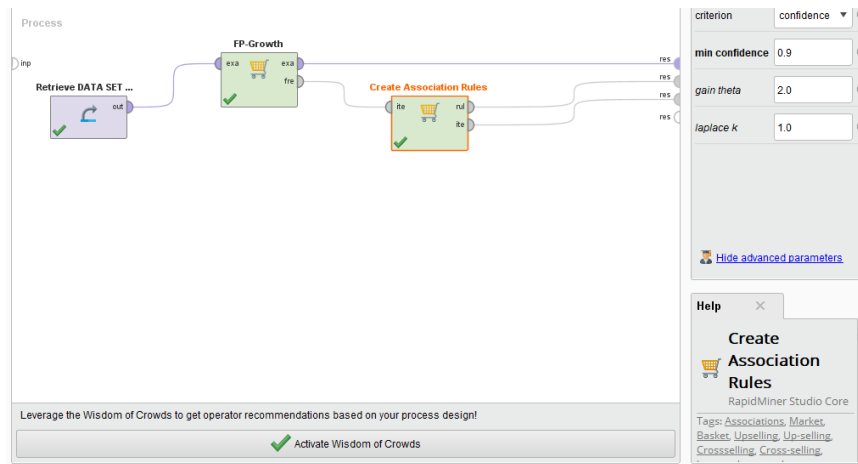
Data transformation dilakukan dengan memberikan inisialisasi terhadap data yang memiliki nilai akan disesuaikan type data yang dibutuhkan pada Algoritma FP Growth. Data Tersebut dapat dilihat pada gambar berikut ini :



Gambar 3. Transformasi

3.4 Data Modeling

Implementasi model algoritma knn sesuai dengan konsep Data Mining maka peneliti dapat di gambarkan model sebagai berikut :



Gambar 4. Model awal.

Berdasarkan Gambar diatas menjelaskan bahwa proses model algoritma membutuhkan tools rapidminer dengan operator Retrive untuk mewardahi dataset, kemudian gunakan operator FP-Growth untuk melakukan mode algoritma dan Set Asiation rule untuk mengetahui pola.

3.5 Evaluation

Berdasarkan hasil confident dengan asosiasi rule menjalaskan bahwa nilai confidence 1 (satu) artinya konsumen membeli item dan diprediksi akan membeli 1(satu) item lagi dengan item yang berbeda hal tersebut tertuang dala data berikut ini.

1. [Udinese] --> [Sassulo] (confidence: 1.000)
2. [Atalanta] --> [Napoli] (confidence: 1.000)
3. [Napoli] --> [Atalanta] (confidence: 1.000)

Dari data diatas mendapatkan pola penjualan, sehingga rekomendasi tersebut dapat dijadikan sebagai bahan pemebelian ulan oleh pemiliki toko.

Berdasarkan hasil confident dengan asosiasi rule menjalaskan bahwa nilai confidence 1 (satu) artinya konsumen membeli 2 (dua) item dan diprediksi akan membeli 1(satu) item lagi dengan item yang berbeda hal tersebut tertuang data data berikut ini.

1. [Bologna, Udinese] --> [Sassulo] (confidence: 1.000)
2. [Spezia, Udinese] --> [Sassulo] (confidence: 1.000)
3. [Fiorentina, Udinese] --> [Sassulo] (confidence: 1.000)
4. [Udinese, Inter] --> [Sassulo] (confidence: 1.000)
5. [Udinese, Empoli] --> [Sassulo] (confidence: 1.000)
6. [Udinese, Torino] --> [Sassulo] (confidence: 1.000)
7. [Udinese, Roma] --> [Sassulo] (confidence: 1.000)
8. [Udinese, Atalanta] --> [Sassulo] (confidence: 1.000)
9. [Udinese, Napoli] --> [Sassulo] (confidence: 1.000)
10. [Sassulo, Atalanta] --> [Napoli] (confidence: 1.000)
11. [Sassulo, Napoli] --> [Atalanta] (confidence: 1.000)
12. [Fiorentina, Empoli] --> [Bologna] (confidence: 1.000)
13. [Bologna, Torino] --> [Fiorentina] (confidence: 1.000)
14. [Fiorentina, Roma] --> [Bologna] (confidence: 1.000)
15. [Bologna, Atalanta] --> [Napoli] (confidence: 1.000)
16. [Bologna, Napoli] --> [Atalanta] (confidence: 1.000)
17. [Udinese, Atalanta] --> [Napoli] (confidence: 1.000)
18. [Udinese, Napoli] --> [Atalanta] (confidence: 1.000)
19. [Inter, Atalanta] --> [Napoli] (confidence: 1.000)
20. [Inter, Napoli] --> [Atalanta] (confidence: 1.000)

Dari data diatas mendapatkan pola penjualan, sehingga rekomendasi tersebut dapat dijadikan sebagai bahan pemebelian ulan oleh pemiliki toko.

Berdasarkan hasil confident dengan asosiasi rule menjelaskan bahwa nilai confidence 1 (satu) artinya konsumen membeli 3 (tiga) item dan diprediksi akan membeli 1(satu) item lagi dengan item yang berbeda hal tersebut tertuang data data berikut ini.

1. [Bologna, Spezia, Udinese] --> [Sassulo] (confidence: 1.000)
2. [Bologna, Fiorentina, Udinese] --> [Sassulo] (confidence: 1.000)
3. [Sassulo, Fiorentina, Empoli] --> [Bologna] (confidence: 1.000)
4. [Spezia, Fiorentina, Udinese] --> [Sassulo] (confidence: 1.000)
5. [Spezia, Udinese, Inter] --> [Sassulo] (confidence: 1.000)
6. [Sassulo, Udinese, Atalanta] --> [Napoli] (confidence: 1.000)
7. [Sassulo, Udinese, Napoli] --> [Atalanta] (confidence: 1.000)
8. [Udinese, Atalanta, Napoli] --> [Sassulo] (confidence: 1.000)
9. [Bologna, Spezia, Empoli] --> [Fiorentina] (confidence: 1.000)
10. [Spezia, Fiorentina, Empoli] --> [Bologna] (confidence: 1.000)
11. [Bologna, Spezia, Fiorentina, Udinese] --> [Sassulo] (confidence: 1.000)

Dari data di atas mendapatkan pola penjualan, sehingga rekomendasi tersebut dapat dijadikan sebagai bahan pemebelian ulan oleh pemiliki toko.

Berdasarkan hasil pengelolaan support didapatkan bahwa konsumen dengan membeli 1 item jersey akan membeli kembali 1 item jersey yang berbeda sebanyak 15 Pola. Konsumen dengan membeli 2 item jersey akan membeli kembali 1 item jersey yang berbeda sebanyak 46 Pola. Konsumen dengan membeli 3 item jersey akan membeli kembali 1 item jersey yang berbeda sebanyak 37 Pola. Konsumen dengan membeli 4 item jersey akan membeli kembali 1 item jersey yang berbeda sebanyak 10 Pola. Sehingga data pol tersebut menjadi rekomendasi owner untuk melakukan pembelian ulang.

4. KESIMPULAN

Hasil penelitian yang telah dilakukan penulis dapat menarik kesimpulan bahwa hasil pengelolaan support didapatkan bahwa konsumen dengan membeli 1 item jersey akan membeli kembali 1 item jersey yang berbeda sebanyak 15 Pola. Konsumen dengan membeli 2 item jersey akan membeli kembali 1 item jersey yang berbeda sebanyak 46 Pola. Konsumen dengan membeli 3 item jersey akan membeli kembali 1 item jersey yang berbeda sebanyak 37 Pola. Konsumen dengan membeli 4 item jersey akan membeli kembali 1 item jersey yang berbeda sebanyak 10 Pola. Sehingga data pol tersebut menjadi rekomendasi owner untuk melakukan pembelian ulang.

REFERENCES

- [1] Br Ginting, A. O. (2021). Penerapan Data Mining Korelasi Penjualan Spare Part Mobil Menggunakan Metode Algoritma Apriori (Studi Kasus: CV. Citra Kencana Mobil). *Journal of Information and Technology*, 1(2),83–90.<https://doi.org/10.32938/jitu.v1i2.1472>.
- [2] Firmasyah, D., Nendriawan, M. R., Studi, P., Informasi, S., Tinggi, S., Informatika, M., & Ilmu, D. (2020). *IMPLEMENTASI ALGORITMA APRIORI PADA DATA PENJUALAN FRISIAN FLAG DI PT. TRIYANTO SUKSES MANDIRI*.
- [3] Frank, Eibe, et al. (2004). *The WEKA Data Mining Software: An Update. Department of Computer Science. New Zealand. University of Waikato*Hamilton.
- [4] Hartanti, D., Nofikasari, I., Yusuf Hermawan, M. A., & Handayani, D. (2021). Penerapan Association Rule Menggunakan Apriori Untuk Rekomendasi Produksi Roti. *Bianglala Informatika*, 9(1), 17–23. <https://doi.org/10.31294/bi.v9i1.9941>
- [5] Kristanto, A. (2019). Tinggi Berdasarkan Prestasi Dengan Menggunakan Metode Iterative Dichotomizer 3 (Id3) (Studi Kasus : Fti Uksw). *Simetris*, 10(2), 433–444.
- [6] Kusriani. (2007). *Strategi Perancangan dan Pengelolaan Basis Data*. CV Andi Offsite.
- [7] Larose, D. T. (2005). *Discovering Knowledge In Data: An Introduction to Data mining*. JohnWiley& Sons. Inc.
- [8] Pratama, I. W., Hafiz, A., Informatika, J. M., & Informatika, J. M. (2019). Implementasi Data Mining Untuk Menentukan Trend Penjualan Cetakan Sablon Pada Fatih Clothing Di Bandar. *Cendekia*, XVIII(1), 326–330.
- [9] Sigit Riyadi. (2017). Penerapan Metode Naive Bayes dalam Pengklasifikasi Trafik Jaringan. *Smatika Jurnal*, 07(1), 20–23.