

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Sebelumnya

Penelitian dilakukan untuk menentukan suatu tata letak pada UI (*User Interface*) pada produk yang ditampilkan pada *e-commerce*. Tidak sedikit pula penelitian mampu memberikan hasil yang cocok untuk diterapkan sehingga dapat digunakan dalam toko atau *online* retail. Dari beberapa penelitian sebelumnya memiliki hasil yang dapat dimanfaatkan untuk memprediksi barang telah dibeli oleh konsumen. Berikut penelitian terdahulu yang dapat digunakan menjadi bahan referensi dalam menyusun penelitian, sehingga dapat memberikan gambaran mengenai sebuah proses dan hasil yang relevan dengan penelitian yang akan dilakukan.

Penelitian sebelumnya yang telah dilakukan Ramot Simangunsong dalam penerapan Market Basket Analysis di CV. Okta. Penulis menggunakan algoritme Apriori dalam menganalisis data penelitian. Dalam penelitian mendapatkan hasil kombinasi produk yang diambil untuk menentukan tata letak dari hasil data transaksi penjualan pada November 2016. Sehingga dapat ditarik kesimpulan bahwa seseorang membeli Mizone maka akan membeli Coolant [10].

Penelitian selanjutnya dilakukan oleh Adila Safitri, Hindriyanto Dwi Purnomo dalam penerapan Market Basket Analysis dengan menggunakan 1000 *record* data bisnis percetakan digital. Dari hasil penelitian mendapatkan 2 item yang berkorelasi (item) tinggi, sehingga dapat merepresentasikan kekuatan pada bisnis tersebut. Mewakili peminat item bagi pelanggan. Lalu dari 2 item tersebut dapat mewakili karakteristik pelanggan. Sehingga dapat meningkatkan layanan dan pemasaran [9].

Pada penelitian selanjutnya yang dilakukan oleh Yudhistyra dan kawan-kawan. Dalam melakukan *Big data analysis* dengan menggunakan *association rule* yang bertujuan untuk menentukan aturan yang memberikan pola hubungan antar item yang jelas, sehingga tim pemasaran memiliki antusias untuk membuat iklan

dalam mempromosikan dalam setiap musimnya, seperti (musim Idul Fitri untuk umat Muslim dan Natal untuk umat Kristiani) [11].

Pada penelitian selanjutnya yang dilakukan oleh Imroatun dan Adhie Thyo. Dalam menentukan *association rules* untuk market basket analysis di TB.Menara menggunakan 1314 data transaksi dan 221 item barang yang paling laku di jual. Dari hasil penelitian yang didapatkan terdapat 1 item yang laku terjual yakni semen holcim yang nilai *support* 48%. Dan item 2 yang laku terjual keramik dan semen dengan nilai *confidence* 89%. Sehingga hasil penelitian dapat membantu mengatasi stok kosongan barang [12].

Pada penelitian selanjutnya yang telah dilakukan Moh.Sholik, Abu Salam. Dalam membangun sistem penjualan yang ada di toko atau perusahaan menggunakan metode asosiasi dengan algoritme Apriori dengan aturan asosiatif kombinasi antar item. Sehingga hasil *association* tersebut mampu digunakan dalam menentukan stok barang yang perlu ada di *supplier*. Dari hasil pengujian analisis juga dapat membantu pihak *E-commerce* dalam mengatur tata letak barang yang dipasarkan [13].

Pada penelitian selanjutnya yang dilakukan oleh Iqbal dan Sarker. Menjelaskan bahwa algoritme Apriori lebih cocok untuk dataset yang besar dibandingkan dengan algoritme Eclat, Lalu algoritme Apriori juga menghasilkan set item data yang sering muncul, sehingga lebih mudah dalam menentukan prediksi belanja *customer* [14].

Tabel 2.1 Ringkasan penelitian sebelumnya

No	Judul	Penulis	Studi Kasus	Metode	Hasil	Perbedaan dengan penelitian yang dilakukan
1.	Pengaruh <i>User Interface</i> dan Variasi Produk terhadap Minat Beli Konsumen. 2023. [22]	Farhan Saputra, Nurul Khaira, Raihan Saputra	Cara penjualan melalui minat beli konsumen marketplace atau <i>E-commerce</i> terhadap <i>User Interface</i> dan Variasi produk	<i>Study Literatur</i>	<i>User Interface</i> Dan Variasi Produk berpengaruh terhadap Minat Beli Konsumen	Penelitian ini fokus dengan study literatur pada pengaruh <i>User Interface</i> dan Variasi Produk.
2.	Penerapan Algoritme Eclat dan Apriori pada <i>Data Mining Market Baset Analisis Penjualan</i> . 2022 [7]	Nirwana Hendrastuty, Ahmad Ari, Aldino, Fikri Hamidy, A. Ferico	Strategi bisnis dengan mencermati pola-pola pembelian pada 212 <i>Mart</i>	Algoritme Apriori dan Eclat	algoritme Eclat menghasilkan rule 86 item dengan waktu 0.01s, algoritme Apriori menghasilkan rule sebanyak 61 rule dalam waktu 0,00s. Didapatkan oleh kedua algoritme terdapat rule yang sama yaitu relasi antara item Frisian Flag Hilo French Vanilla 225 MI Dan Frisian Flag Low Fat Belgian Chocolate 250ml. Algoritme Eclat dalam urutan ke 5 sedangkan algoritme Apriori termasuk dalam urutan ke 3 dengan nilai support 0.008230453	Pada Penelitian sebelumnya menggunakan 2 algoritme Apriori dan Eclat. Menggunakan Bahasa pemrograman R.
3.	<i>Market basket analysis for Fresh Products location improvement: A case study of E-Commerce Business Warehouse</i> . 2022 [15]	Pattraporn Jirapatsil Naragain Phumchusri	Menganalisis 2366 data transaksi dari sebuah perusahaan produk segar secara elektronik/ <i>E-commerce</i> untuk	Algoritme Apriori	Jarak rata-rata per pesanan berkurang dari 31,8 meter menjadi 14,5 meter, atau berkurang 54,4%.	Pada penelitian sebelumnya menganalisis transaksi pembelian produk segar untuk menghitung jarak tempuh <i>customer</i> dalam pembelian

No	Judul	Penulis	Studi Kasus	Metode	Hasil	Perbedaan dengan penelitian yang dilakukan
			mengatur rata-rata pesanan			
4.	<i>Student Behavior Data Analysis Based on Association Rule Mining.</i> 2022. [17]	Tengfei Wang, Baorong Xiao, Weixiao Ma	Penerapan <i>association rule mining</i> dalam analisis perilaku mahasiswa. Untuk mengatasi masalah pemindaian asli dan kelebihan iterasi	<i>Association rule</i>	Hasil analisis kasus menunjukkan bahwa, algoritme <i>association rule</i> yang ditingkatkan secara efektif mampu meningkatkan efisiensi penambangan data.	Penelitian sebelumnya menggunakan <i>student behavior</i> untuk mendukung promosi smart campus. Sehingga dapat membantu <i>management</i> kampus
5.	Market basket analysis Untuk Menganalisis Pola Perilaku Konsumen Wahdana Mall. 2021 [18]	Erni Rouza, Nurhayati, Mona	Model perilaku konsumen yang dapat dianalisis adalah pola jenis produk yang sering dibeli oleh konsumen berdasarkan <i>list</i> belanja, sehingga diperoleh <i>rule-rule</i> dari hubungan item-item produk yang dibeli oleh konsumen secara bersamaan.	<i>Market basket analysis</i>	Metode Market Basket Analysis khususnya <i>Association Rule</i> mampu menghasilkan informasi pengetahuan atau <i>knowledge</i> baru yang nantinya bisa digunakan dalam proses pengambilan keputusan untuk penentuan strategi pemasaran dan penjualan.	Penelitian sebelumnya Menggunakan pengembangan dengan algoritme Apriori tetapi kurangnya kriteria dalam sataset yang digunakan, sehingga pola <i>rules</i> yang didapatkan kurang maksimal
6.	Analisa Data Penjualan Handphone dan Elektronik Menggunakan Algoritme Apriori 2021. [19]	Fadhila Shely Amalis, Setiawansya, Dedi Darwis	Menganalisis pola belanja pada transaksi CV Rey Gasendra	Algoritme Apriori	membandingkan perhitungan manual dengan perhitungan yang dilakukan dengan rapidminer menggunakan 100 data sampel memperoleh hasil kesimpulan yang sama yaitu 15 <i>rules</i> .	Penelitian sebelumnya Menggunakan Algoritme Apriori dengan metode rapidminer
7.	Analisa data transaksi penjualan barang menggunakan algoritme Apriori dan FP-Growth. 2020 [5]	Harianto, Hadryan Eddy	Mengembangkan algoritme Apriori dengan perbandingan FP-Growth	Algoritme Apriori dan FP-Growth	Algoritme Apriori membutuhkan waktu 0.0424 detik atau 39% sedangkan algoritme FP-Growth	Pada penelitian sebelumnya menggunakan dataset dengan file transaksi yang terpisah-pisah. Sehingga membutuhkan

No	Judul	Penulis	Studi Kasus	Metode	Hasil	Perbedaan dengan penelitian yang dilakukan
			menggunakan data transaksi retail 2021		membutuhkan waktu 0.0650 detik atau 61%	waktu lebih lama di saat data <i>preprocessing</i>
8.	Implementasi Data Mining Menentukan Rekomendasi Penempatan Buku Berdasarkan Pola Peminjaman Dengan Menggunakan <i>Association Rule</i> . 2020	Yusuf Nawawi, Imam Suharjo, S.T., M.Eng.	Keterhubungan antara item buku yang dipinjam diperiksa dengan analisis data peminjaman buku perpustakaan dari 2014 sampai dengan Maret 2019	<i>Association rule</i>	Jumlah data 2225 <i>Support</i> dan <i>confidence</i> 0.01 dengan pembatasan 50 aturan asosiasi, dengan <i>support</i> x <i>confidence</i> 8.17%.	Pada penelitian sebelumnya Menganalisis pola peminjaman untuk rekomendasi penempatan kategori buku dengan <i>association rules</i> pada Perpustakaan[16]
9.	Algoritme FP-Growth Untuk Menganalisis Pola Pembelian Oleh-Oleh 2020 [20]	Yola Permata Bunda	Menentukan strategi pemasaran dalam meningkatkan penjualan di Pusat Oleh-Oleh Ummi Auifa Hakim.	Algoritme Fp-Growth	jika membeli Cincang Kuning 250 gr, maka akan membeli Sj.Balado Merah 250 gr memiliki nilai <i>support</i> 20 dan nilai <i>Confidence</i> 100%, kemudian apabila membeli Ganepo 250gr, mungkin membeli Sj.Balado Merah dengan nilai <i>support</i> 20% dan nilai <i>Confidence</i> 100%	Penelitian sebelumnya menggunakan Algoritme Fp-Growth
10.	<i>An Improved Apriori Algorithm Based on Support Weight Matrix for Data Mining in Transaction Dataset</i> . 2020. [21]	Li-na Sun	Melakukan pembuktian pada proses data mining dengan algoritme Apriori untuk membangun matrik 0-1 dengan memindai basis data transaksi	Algoritme Apriori	Algoritme lebih baik mempersingkat waktu berjalan dan mengurangi kebutuhan memori dan jumlah operasi I/O. pada dukungan untuk item langka cenderung meningkat, sedangkan dukungan untuk item lain sedikit menurun, sehingga item yang tersembunyi dan berharga dapat secara efektif	Penelitian sebelumnya melakukan pembuktian performa algoritme Apriori akan semakin baik lagi dengan pembobotan pada matrik karena mampu mempersingkat waktu dan mengurangi memori

No	Judul	Penulis	Studi Kasus	Metode	Hasil	Perbedaan dengan penelitian yang dilakukan
					dan barang berharga dapat diekstraksi secara efektif	
11.	Implementasi Algoritme Apriori untuk Mencari Asosiasi Barang yang Dijual di <i>E-commerce</i> OrderMas 2018 [13].	Moh.Sholik, Abu Salam	Menganalisis data transaksi penjualan untuk menyelesaikan masalah mengenai penyediaan stok barang oleh <i>supplier</i> OrderMas	Algoritm Apriori	Dengan 324 data mendapatkan nilai minimal support 50% dan minimal <i>confidence</i> 50% dan mendapatkan jumlah <i>rule</i> sebanyak 10 <i>rules</i> .	Pada penelitian sebelumnya menggunakan nilai <i>minimum support</i> dan <i>minimum confidence</i> yang sama sehingga hasil <i>rule</i> yang didapatkan kurang bervariasi

Tabel yang menyajikan penelitian yang telah dilakukan sebelumnya. Data yang tercantum pada jurnal memberikan wawasan untuk penelitian yang akan dilakukan. Mencakup berbagai topik yang berhubungan dengan subjek penelitian yang akan dilakukan. Dari penelitian sebelumnya memberikan gambaran yang lebih luas tentang perkembangan Market Basket Analysis. Metode memberikan gambaran hasil yang telah dilakukan. Sehingga membuka peluang baru untuk penelitian di masa depan.

Pada penelitian yang telah dilakukan oleh Farhan Saputra dkk., membandingkan 2 algoritme Apriori dan Eclat. Dari kedua algoritme ternyata ditemukan bahwa algoritme Apriori mendapatkan waktu yang lebih cepat dibanding algoritme Eclat, dengan asil *rules* yang berbeda. Penelitian yang telah dilakukan Nirwana Hendrastuty dkk. menjelaskan mengenai penggunaan algoritme Apriori yang dilakukan dalam e-commerce. Tetapi didalamnya tidak dijelaskan bagaimana aturan yang digunakan untuk membuat tata letak itu dibuat secara statis atau dinamis sesuai dengan keadaan konsumen. Pada penelitian Lina-Sun telah melakukan pengujian untuk algoritme Apriori ternyata bagus apabila terus dikembangkan lagi dengan matrik 0-1 atau (true/false). Sehingga hal ini menjadi tambahan wawasan untuk mengambil penelitian Market Basket Analysis untuk menentukan saling keterhubungan antar produk yang

digunakan oleh pelanggan. Pada penelitian sebelumnya Moh.Sholik menjelaskan mengenai keterkaitan User Interface dengan variasi produk yang disajikan, itu sangat berpengaruh pada e-commerce Maka dalam penelitian ini menggunakan data yang lebih besar dari penelitian sebelumnya. Sehingga nantinya mendapatkan analisis produk sesuai dengan preferensi pelanggan dengan Market Basket Analysis yang dinamis dalam tampilan variasi produk.

2.2 Landasan Teori

2.2.1 *Customer Behavior*

Perilaku konsumen (*customer behavior*) yaitu sebagai sebuah proses dan tindakan konsumen yang ikut terlibat ketika mencari, memilih, menggunakan, mengevaluasi, dan membeli berbagai aneka produk dan jasa yang digunakan untuk kepuasan konsumen [23]. Perilaku konsumen melibatkan banyak hal untuk mencapai kepuasan konsumen atau hal yang diinginkan konsumen. Perilaku pelanggan yang melibatkan beberapa faktor dalam mempengaruhi keputusan dalam pembelian suatu produk. Beberapa faktor mencakup emosional, mental dan perilaku konsumen yang lainnya [24]. Memahami *customer behavior* sangat penting untuk bisnis untuk membantu mencari strategi pemasaran yang efektif.

Produk yang dibeli oleh konsumen seringkali akan memiliki pengaruh yang signifikan dalam keputusan pembelian konsumen yang lainnya [25]. Produk dengan penjualan yang terlaris mampu meningkatkan penjualan. Produk yang berhasil mendominasi pasar dan menjadi favorit konsumen memiliki daya tarik yang kuat untuk mempengaruhi keputusan pembelian konsumen lainnya. Secara langsung memacu peningkatan penjualan. Ketika banyak konsumen memilih produk tertentu pada suatu *e-commerce*, produk tersebut secara otomatis mendapatkan reputasi yang positif. Rekomendasi dari konsumen yang puas dengan suatu produk cenderung memberikan dampak yang signifikan terhadap keputusan pembelian orang lain. Hal ini membuat produk tersebut menjadi semacam *seal of approval* dalam komunitas konsumen.

Saat ini banyak *e-commerce* untuk menarik konsumen dengan produknya dengan melakukan perubahan tata letak barang yang berdekatan atau sering muncul dalam *e-commerce*, Produk yang dibeli sering kali membuat pelanggan yang lain menjadikannya referensi dalam berbelanja. Sehingga pola pembelian ini saling berhubungan antar produk yang dibeli oleh konsumen. Hal ini mampu di jadikan prediksi tata letak produk dengan preferensi pola pelanggan. Sebagai solusi untuk meningkatkan transaksi.

2.2.2 User Interface

User interface adalah tampilan visual sebuah produk yang menjembatani sistem dengan pengguna (*user*) [26]. Pengembangan perangkat lunak *user interface* mengacu pada cara pengguna berinteraksi dengan sistem atau aplikasi melalui elemen visual, masukan dan tindakan yang tersedia. UI sebagai jembatan komunikasi antara pengguna dan program, memungkinkan pengguna dengan mudah dan efektif mengontrol dan memanfaatkan fungsi sistem. Tujuan dari sistem untuk membantu merancang tata letak produk untuk membantu proses penjualan yang nantinya akan membantu proses promosi dan meningkatkan transaksi konsumen [27]. Produk yang ditampilkan akan menjadi referensi konsumen agar membeli produk yang lain. Sehingga pembelian pada *e-commerce* mampu meningkat. Tampilan *user interface* berpengaruh terhadap minat beli konsumen. *User interface* merupakan alat komunikasi dengan konsumen untuk mempengaruhi konsumen dengan membagi informasi terkait produk sebagai langkah untuk promosi [28].

2.2.3 Data Mining

Data mining merupakan salah satu teknik dalam proses mempekerjakan satu atau lebih teknik pada pembelajaran komputer untuk menganalisis dan mengekstrak data secara otomatis [29]. Proses dalam menjalankan satu atau beberapa data dalam eksekusi Machine Learning secara otomatis dengan tujuan melakukan analisis dan melakukan ekstraksi dalam *knowledge* yang informasinya secara tersirat. *Data mining* merupakan langkah dalam mencari pengetahuan atau informasi dari *database Knowledge Discovery* [30]. Tujuannya untuk memahami pengetahuan yang tersembunyi dari dataset dengan analisis statistik konvensional yang secara umum mempunyai makna yang tersirat. Terdapat 6 tahapan dalam *data mining* [30], yaitu :

- a. *Data Integration* adalah tahapan dimana data yang lebih dari satu *database* digabungkan menjadi satu database utama dalam konteks *data mining*. Dalam *data mining*, terkadang diperlukan penggunaan lebih dari satu *database* untuk memenuhi kebutuhan analisis data. Proses ini melibatkan penggabungan data

dari berbagai sumber dengan format yang berbeda menjadi satu kesatuan yang terstruktur dan konsisten. Tujuan dari *data integration* adalah menciptakan *database* utama yang mencakup semua data yang diperlukan untuk analisis *data mining* yang komprehensif dan efektif.

- b. *Data Selection* adalah proses pemilihan atribut-atribut suatu *database* dan hapus atribut suatu data yang tidak digunakan. Tujuannya adalah untuk menyaring dan memilih atribut-atribut yang relevan dan penting dalam konteks analisis data. Dalam proses ini, atribut-atribut yang tidak terpakai atau tidak memberikan kontribusi signifikan terhadap tujuan analisis data akan diidentifikasi dan dihapus. Dengan melakukan pemilihan atribut yang tepat, dapat mengurangi kompleksitas data, mempercepat proses analisis dan meningkatkan kualitas hasil.
- c. *Data Cleaning* adalah suatu tahapan dimana dilakukan upaya untuk membersihkan data dengan tujuan menghilangkan data yang tidak perlu atau tidak relevan dalam konteks *data mining*. Dalam proses ini, data yang tidak perlu atau tidak relevan akan diidentifikasi kemudian dihapus sehingga data hanya tersisa menjadi lebih berkualitas dan dapat memberikan informasi yang lebih bernilai dalam analisis data.
- d. *Data Transformation* adalah suatu langkah dalam pemrosesan untuk mengubah data dalam format dengan ketentuan yang sesuai dengan kebutuhan analisis atau pengolahan lebih lanjut. Dalam tahap ini, data dapat mengalami perubahan struktur, skala, atau representasi agar dapat dimanfaatkan dengan lebih efektif. Tujuan utama dari *data transformation* adalah menghasilkan data yang lebih terstruktur, terstandarisasi, atau terformat sehingga memudahkan analisis, integrasi, atau pemodelan data.
- e. *Process Mining* adalah tahapan utama untuk penerapan metode yang bertujuan menghasilkan informasi atau *knowledge* baru yang berharga yang sebelumnya tersembunyi. Dalam proses ini, data *log* yang terkait dengan aktivitas bisnis atau proses operasional dieksplorasi dan dianalisis secara sistematis. Melalui teknik analisis khusus, seperti *discovery*, *conformance*, dan *enhancement*, informasi yang berharga dapat diungkapkan tentang bagaimana proses

sebenarnya dilakukan, mengidentifikasi penyimpangan atau masalah, dan mengidentifikasi peluang untuk meningkatkan efisiensi dan kualitas proses.

- f. *Pattern Evaluation* adalah tahap dalam proses *mining* di mana hasil *mining* yang telah dilakukan dievaluasi dengan mengidentifikasi pola-pola yang muncul. Pada tahap ini, pola-pola yang relevan dan menarik yang telah ditemukan dalam *data mining* dianalisis lebih lanjut untuk memahami signifikansi, kegunaan, dan dampaknya terhadap tujuan analisis.
- g. *Knowledge Presentation* adalah tahap di mana informasi mengenai metode yang digunakan dan kesimpulan dari proses *mining* disajikan secara jelas dan komprehensif. Pada tahap ini, hasil analisis yang telah dilakukan dan pengetahuan yang ditemukan dalam proses *mining* direpresentasikan dalam bentuk yang dapat dimengerti oleh pengguna atau pemangku kepentingan.

2.2.4 Market Basket Analysis

Market basket analysis adalah proses analisis perilaku dengan khusus atau spesifik dari kelompok tertentu. Analisis keranjang pasar adalah proses menganalisis suatu kebiasaan transaksi pelanggan dengan menemukan hubungan antara produk yang berbeda dalam keranjang belanja pelanggan [31]. Biasanya digunakan sebagai langkah awal untuk mengetahui nilai transaksi ketika kita tidak tahu pola secara khusus dari yang dicari [32]. Analisis keranjang pasar dapat digunakan untuk banyak membuat keputusan, terutama tata letak produk [15]. Sehingga menimbulkan identifikasi keterkaitan pada sekelompok produk, item maupun kategori. Konsumen yang sering membeli produk memiliki kaitan dalam membentuk hubungan jika-maka yang disebut dengan *rule if-then* [12].

Market basket analysis salah satu metode yang mampu dalam memberikan solusi dari rekomendasi kombinasi pembelian kategori produk yang berkaitan dengan menggunakan hubungan *association rule*[4]. Analisis yang dilakukan dalam *Market Basket Analysis* salah satunya dengan cara melihat perilaku konsumen yang melakukan pembelian produk. Dari proses pembelian produk maka akan muncul hubungan antara produk yang dibeli oleh *customer*.

2.2.4.1 Association Rules

Association rule merupakan salah satu proses dalam teknik *data mining* yang dapat digunakan untuk menentukan sejumlah aturan atau kombinasi rekomendasi dari item atau *frequent itemsets* [4]. *Association rules* dapat mendeteksi berbagai fitur dalam dataset. Aturan ini dapat dilakukan untuk mencari hubungan data dengan *customer behavior* dengan informasi produk untuk fitur-fitur yang menjadi penentu tata letak atau *layout* suatu produk [33]. Maka konsep pada *association rules* untuk menemukan pola produk yang sering muncul dalam peristiwa transaksi *customer behavior* mendukung teknik untuk rekomendasi produk dengan *layout* yang dihasilkan [34]. Dalam menentukan pola itemset dengan *association rules* menggunakan ukuran ketertarikan (*interestingness measure*) dengan batasan yang telah ditentukan pada saat dihasilkan dari perhitungan pengolahan data [35]. Kombinasi batas yang dimiliki untuk *interesting* yaitu kombinasi suatu itemset apabila telah bisa memenuhi *minimum support* dan nilai *minimum confidence* dari *rules* yang sudah dibuat sebelumnya tersebut [36].

a. Analisis Pola Frekuensi Tinggi

Nilai *support* adalah nilai penunjang antara item atau produk [30]. Jumlah item dengan semua isi data transaksi. Nilai yang diperoleh yang digunakan untuk mengukur seberapa sering item tingkat kesamaan yang dimiliki dalam transaksi dari keseluruhan transaksi [37]. Semakin tinggi nilai *minimum support* yang di dapat membuat *rule* yang dihasilkan menjadi semakin sedikit, begitupun sebaliknya apabila nilai *minimum support* terlalu kecil maka *rule* yang didapatkan akan semakin banyak. Nilai *support* dapat dihitung dengan rumus perhitungan [38]:

$$Support(A, B) = \sum \frac{\text{Transaksi mengandung A dan B}}{\text{Transaksi}} \quad (2.1)$$

b. Pembentukan Association Rules

Nilai *minimum confidence* menentukan proses pengukuran hubungan antar item dengan satu item yang lain. Nilai *confidence* bentuk kepastian dari item yang terbentuk. Semakin tinggi nilai *minimum confidence* yang dicapai maka

semakin mendapatkan hasil rule yang sedikit. Untuk mencari nilai *minimum confidence* dapat dihitung dengan rumus:

$$Confidence(A, B) = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi mengandung A}} \quad (2.2)$$

c. Proses Perhitungan *Lift Ratio*

Pengukuran ketepatan untuk hubungan antara *association* dari nilai *minimum support* dan nilai *minimum confidence*. Apakah hasil yang didapatkan dari aturan valid atau tidak valid [39]. Menurut Jiawei Han, Apabila korelasi positif *lift ratio* bernilai >1 maka kemunculan satu item menyiratkan kemunculan yang lainnya, Apabila nilai korelasi <1 menjadi korelasi negatif maka salah satu kemungkinan besar ke tidak adanya hal yang lain, jika nilai $=1$ maka diartikan korelasi netral karena kemunculan kedua item disebut tidak saling bergantung [37]. Untuk mencari nilai *life rasio* menggunakan rumus

$$Lift Ratio = \frac{P(A \cup B)}{P(A) * P(B)} \quad (2.3)$$

Keterangan :

$P(A \cup B)$: Peluang kombinasi kemunculan dari item A dan B

$P(A)$: peluang kemunculan A

$P(B)$: Peluang kemunculan B

2.2.4.2 Algoritme Apriori

Algoritme adalah prosedur beberapa langkah demi langkah perhitungan. Algoritme yang digunakan perhitungan, penalaran otomatis, dan pemrosesan informasi. Pengertian algoritme adalah metode yang secara efektif dinyatakan sebagai rangkaian yang dibatasi oleh beberapa instruksi dijelaskan dengan baik untuk hitung suatu fungsi [40]. Algoritme Apriori adalah salah satu algoritme dalam *association rule*. Aturan asosiasi yang lebih baik untuk mencari relasi keranjang belanja. Himpunan item yang didapatkan dari beberapa aturan untuk mengurangi komputasi [41].

Algoritme Apriori merupakan metode yang mencari suatu pola dari hubungan antara pola yaitu dengan banyak pola dalam sebuah dataset. Algoritme Apriori sendiri sering digunakan saat ada transaksi atau yang bisa juga dengan *market basket*. Dengan algoritme Apriori, *owner* perusahaan dapat menemukan pola dalam pembelian barangnya yang sering dibeli oleh *customer*. Dapat mengetahui pola dalam pembelian barang mereka yang dibeli oleh pelanggan [42]. Cara kerja dari algoritme didasarkan pada pertimbangan memeriksa calon itemset dari hasil frekuensi pada itemset menggunakan *support-based pruning* yang tujuannya menyeleksi itemset yang tidak berpengaruh agar terhapus pada pemilihan minimal *support* [43].

2.2.5 Text Mining

Text Mining adalah langkah analisis teks yang dilakukan oleh komputer secara otomatis untuk menggali informasi yang berkualitas dari suatu rangkaian teks yang terkandung dalam sebuah dokumen [44]. *Text mining* bertujuan untuk mengelola dan mengekstrak informasi dari data *text*. Sehingga tujuan akhir memahami makna dari teks yang diolah. Penggalan data dalam menyelesaikan masalah mengenai kebutuhan suatu informasi dengan teknik *data mining*, Machine Learning, Natural Language Processing dan manajemen pengetahuan. Proses kerja *text mining* mengadopsi dari *data mining* namun perbedaannya pada pola yang digunakan dalam *text mining* diambil dari sekumpulan bahasa alami yang tidak terstruktur sedangkan *data mining* pola yang diambil dari *database* yang terstruktur [45]. Tahapan -tahapan yang bisa dilakukan pada proses *text mining* sebagai berikut:

- a. *Tokenizing* adalah tahap yang digunakan untuk menghapus kalimat yang sama dan membagi kalimat, paragraf atau dokumen menjadi bagian yang lebih kecil yaitu token atau kata terpisah yang berdiri sendiri. Token ini bisa berupa kata, frasa atau karakter tergantung pada kebutuhan. Pemecahan teks menjadi token mempermudah analisis yang memungkinkan untuk mengidentifikasi kata-kata yang digunakan. Proses tokenisasi biasanya melibatkan penghapusan karakter non-alfabet, tanda baca dan langkah-langkah untuk memisahkan kata-kata atau frasa

- b. *Stopwords* adalah kata-kata yang tidak memiliki arti dalam sebuah kalimat, bahasa yang biasanya tidak memberikan kontribusi signifikan terhadap makna suatu teks. Menghilangkan *stopwords* yang terdapat pada kalimat atau paragraf maka proses pengolahan data akan lebih mudah dan tidak akan menimbulkan kesalahpahaman dalam proses analisis. Sehingga langkah penghapusan *stopwords* dilakukan untuk meningkatkan kualitas analisis dan fokus pada kata-kata yang lebih informatif. Mengimplementasikan penghapusan *stopwords*, *text processing* menjadi lebih akurat dan efektif dalam menghasilkan wawasan yang bermakna dari data teks yang kompleks.

2.2.6 Algoritme Naïve Bayes

Naïve Bayes adalah metode klasifikasi yang sederhana berdasarkan teorema bayes yang secara efisien dapat melakukan klasifikasi menggunakan beberapa pelatihan data [46]. *Bayesian classification* adalah pengklasifikasian *statistic* yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu *class*. *Bayesian classification* memiliki nilai akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam *database* dengan data yang besar [47]. Sehingga dalam hal ini Naïve Bayes digunakan untuk mencari probabilitas kecenderungan antar *rules* dari *association rules*. Untuk menentukan hasil *keyword* yang sesuai dengan *association rules*. Metode Naïve Bayes yang digunakan dengan memanfaatkan *library scikit-learn*, dengan menggunakan algoritme Multinomial Naïve Bayes. Penggunaan rumus Multinomial untuk mencari probabilitas.

$$P(C|X) = \frac{P(X|C).P(C)}{P(X)} \quad (2.4)$$

Keterangan :

$P(C|X)$: Probabilitas kelas C diberikan observasi X

$P(X|C)$: Probabilitas observasi X diberikan kelas C

$P(C)$: Probabilitas prior untuk kelas C

$P(X)$: Probabilitas marginal dari observasi X