

BAB II

TINJAUAN PUSTAKA

2.1. Tinjauan Teori

Penelitian yang dilakukan oleh peneliti ini membutuhkan dukungan teoritis agar seluruh rangkaian penelitian memiliki fundamental penelitian yang baik dan benar.

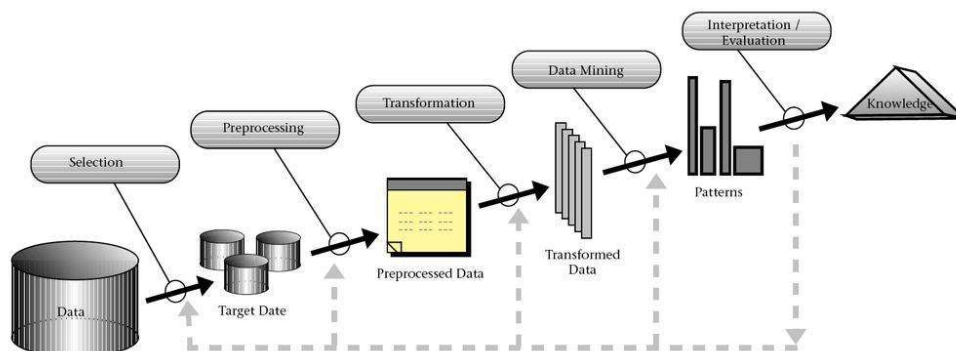
2.1.1. Teori Umum

2.1.1.1. Penemuan Pengetahuan Pada Database (KDD)

Pengelolaan maupun untuk menemukan (*manage/explore*) informasi dari suatu basis data, atau yang lebih sering dikenal dengan istilah KDD, adalah satu proses ataupun metode guna mendapat informasi dari seluruh data tersedia, didalam *database* sendiri akan ditemukan banyak tabel yang memiliki hubungan satu dengan lainnya (berelasi). Dalam rangka membuat beberapa keputusan penting terhadap masalah yang dihadapi, dalam hal ini informasi tersebut (pengetahuan yang dihasilkan) setelah proses mengelola atau menemukan data selesai kemudian digunakan oleh perusahaan (Elisa, 2018)

Melakukan eksplorasi atau menggali data berarti menemukan pola dalam data yang sebelumnya tidak diketahui. Ketika pola yang diciptakan ditemukan, diharapkan hal itu akan menjadi lebih mudah bagi orang atau pihak pembuat keputusan. Data mining adalah bidang ilmu yang membagi penemuan pengetahuan dalam database (KDD). Nampak pada gambar berikut bagaimana data mining dilakukan:

- 1) *Cleaning* atau Pembersihan data, yakni sebuah langkah terukur penghapusan *inconsistent data* (tidak konsisten),
- 2) Integrasi data, yaitu tahap di mana sejumlah basis data digabungkan untuk membentuk basis data yang bisa dikatakan *new database*
- 3) Seleksi Data (*selection*), yaitu proses pemilihan data yang akan digunakan dalam penelitian sebab pentingnya data yang terpilih tersebut,
- 4) *Transforming*, Menghasilkan format data yang disesuaikan untuk digunakan dalam pengolahan data dengan algoritma yang akan digunakan.
- 5) Prosedur penambahan, langkah yang paling penting dalam penambahan data juga dalam menghasilkan informasi/pengetahuan,
- 6) Evaluasi pola, evaluasi informasi/pengetahuan terkait pada pola atau *pattern* terbentuk yang akan dieksekusi oleh pihak tertentu dalam proses pengambilan keputusan (*decision making*).
- 7) Presentasi Pengetahuan, menjelaskan secara visual hasil ekstraksi/penggalan data.



Gambar 2.2 Tahapan dalam KDD

2.1.1.2. Menambang Data (*Data Mining*)

Menambang data (*Data mining*) dikenal pada dunia data science karena kemampuan dalam penggunaan ilmu statistika, pengetahuan pada matematika, pengetahuan pada kecerdasan buatan, serta saat ini pada perkembangan pada teknologi mesin pembelaran (*machine learning*) yang bertujuan mengambil serta mengeksplorasi informasi yang terdapat pada sebuah set basis data (*database*). Penambangan data juga dikatakan sebagai proses atau teknik mengetahui hal yang belum diketahui sebelumnya guna menambahkan nilai tambah (*added value*) (Sigalingging & Harman, 2020).

Penelitian menunjukkan bahwa dengan perkembangan terus-menerus teknologi data mining, sistem tidak hanya membantu pengguna melakukan analisis pengambilan keputusan pemasaran yang ilmiah dan rasional, meningkatkan keberlanjutan dan kecanggihan proses pengambilan keputusan pemasaran, tetapi juga dapat membawa gagasan baru kepada para pengambil keputusan di perusahaan, mendorong peningkatan dan kemajuan berkelanjutan dari sistem tersebut. (Hou et al., 2023)

Teknologi data mining menggunakan berbagai algoritma a priori untuk memastikan aturan asosiasi. Aturan ini menetapkan hubungan antara atribut khusus dalam analisis afinitas, yang juga dikenal sebagai analisis keranjang pasar. Dalam penambangan aturan asosiasi, kombinasi aturan dibentuk pada item. Tujuan dari algoritma apriori adalah untuk membedakan set item yang sering muncul yang dapat diterapkan pada data tertentu, menghasilkan pola yang umumnya teramati selama transaksi. Untuk kemudahan komputasi, perangkat lunak sumber terbuka

RapidMiner digunakan dalam lingkungan data mining. Ini digunakan untuk berbagai tujuan, mencakup data mining, text mining, algoritma pembelajaran mesin, dan pelaksanaan analisis prediktif (Saptadi et al., 2023)

Berdasarkan fungsinya, tugas data mining dapat dikelompokkan ke dalam kelompok berikut (Elisa, 2018):

1. Klasterisasi

Membagi dalam kelompok objek menjadi agar terbagi menjadi kelompok objek yang mirip, diketahui sebuah cluster merupakan keharusan untuk berisi objek yang mirip serta adanya objek antar cluster yang berbeda.

2. Klasifikasi

Melakukan pembagian berdasarkan kelompok objek berdasarkan kelompok.. Perbedaan dengan clustering adalah bahwa classifier ini membutuhkan data yang sudah dilatih (*trained*) sudah diberi label atau tanda menjadi grup maupun kelas.

3. Regresi

Metode ini secara fundamental hampir sama dengan klasifikasi, karena membutuhkan *trained data* yang diberi label atau tanda. Perbedaannya adalah luaran *classifier* bernilai diskrit, tetapi luaran regresi yaitu kontinu. Metode tersebut mengeksplorasi sebuah bentuk pola relasi diantara atribut prediktor dengan dependen, yang mana diketahui atribut dependen merupakan bernilai kontinu juga.

4. Asosiasi

Kelompok data mining ini memiliki tanggung jawab untuk mengetahui karakter secara bersamaan terjadinya. Kelompok data mining Asosiasi digunakan untuk menemukan *rules* dalam membuat ukuran hubungan antara beberapa atribut (dua atau lebih atribut). Aturan akan dibentuk ketika premis (*antecedent*), konsekuensi (*Consequence*), dan keyakinan (*confident*) terkait dengan aturan.

2.1.1.3. Aturan Asosiasi (*Association Rule*)

Analisis asosiasi atau association rule mining adalah teknik data mining untuk menemukan aturan asosiasi antara suatu kombinasi item. *Interestingness measure* yang dapat digunakan dalam data mining adalah (Novianti & Elisa, 2020):

1. Nilai/skor dukungan adalah ukuran seberapa unggul suatu item atau kumpulan item terhadap perdagangan secara keseluruhan.
2. Nilai/skor reliabilitas adalah ukuran yang menunjukkan hubungan antara dua hal secara kondisional (berdasarkan situasi tertentu).

Analisis ini banyak difungsikan sebagai cara dalam menemukan hubungan tersembunyi yang penting antara sekelompok data besar. Hubungan atau relasi tersebut dinyatakan sebagai aturan asosiasi maupun aturan pada elemen.

Analisis asosiasi berguna untuk menemukan hubungan penting antara elemen-elemen dari setiap transaksi, yang dapat menunjukkan kekuatan aturan dalam asosiasi. Untuk menemukan aturan asosiasi yang diinginkan, Anda perlu menemukan nilai dukungan yang ditentukan. Support adalah jumlah elemen dalam setiap transaksi pada kumpulan data. mencari nilai support tersebut dibutuhkan, bisa dicari seluruh rule yang ekuivalen dengan pengurangan support tersebut.

Untuk itu bisa digunakan menjadi proses tahapan dalam mencari nilai *confidence*. *Confidence* akan ditentukan dari nilai *support rule* yang didapatkan pada transaksi.

Apabila nilai skor item set setiap transaksi tidak sering muncul, kandidat yang tidak sesuai dengan nilai support minimum akan langsung dipotong tanpa harus menghitung nilai *confidence*-nya. Strategi umum yang digunakan oleh banyak algoritma penambangan aturan asosiasi adalah membagi masalah menjadi dua tugas utama sebagai berikut;

1. Pengembangan freq itemset: Tujuannya adalah untuk memastikan bahwa semua itemset memenuhi ambang batas minimum dukungan. Itemset frekuen, atau itemset yang sering muncul disebut itemset *frequent*.
2. Aturan yang dihasilkan bertujuan dalam mengambil *rule* dimana *confidence* yang tinggi dari itemset frekuent yang didapatkan dari proses yang terjadi sebelum itu, kemudian hal ini disebut aturan skor tinggi (*strong rules*)

2.1.2. Teori Khusus

2.1.2.1. Algoritma Apriori

Dalam dunia eksplorasi data terutama pada data mining ada beberapa metode penggalian data, termasuk algoritme apriori, yang mendeteksi asosiasi, tren, dan aturan asosiasi di antara item yang dibeli. Algoritma ini bekerja dengan menghasilkan pola "jika - maka". Informasi aturan asosiasi dari transaksi pembelian konsumen dapat digunakan untuk membantu menyediakan informasi stok barang yang sesuai. Stok dapat diseimbangkan dengan mempertimbangkan barang-barang yang sering dibeli secara bersamaan. Strategi promosi dapat dibuat berdasarkan kombinasi barang yang sering dibeli untuk meningkatkan penjualan. Barang-

barang juga dapat diatur di minimarket berdasarkan barang yang sering dibeli. *Algoritma association rule mining* lainnya bisa difungsikan dalam mengekstrak aturan asosiasi pada data yang ada didalam *algoritma Predictive Apriori* serta *algoritma Tertius*. Apriori dipilih lebih cenderung dari algoritma Predictive ataupun tertius disebabkan oleh algoritma Apriori menghasilkan aturan asosiasi yang lebih baik (Alma et al., 2020)

Algoritma data mining dengan jenis Apriori merupakan algoritma pengambilan data dimana aturan asosiasi digunakan dalam menentukan bagaimana relasi antara kombinasi item yang terdapat didalam basis data. Algoritma Apriori menemukan kombinasi entri dalam *database* yang memenuhi nilai dukungan minimum untuk analisis pola frekuensi tinggi. Nilai support suatu item ditentukan dengan menggunakan rumus berikut untuk menentukan nilai support

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

Nilai Support dari 2 item diperoleh dengan menggunakan rumus:

$$\text{Support (A, B)} = P(A \cap B)$$

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

Rumus 2.1 Perhitungan skor support apriori

Pembentukan aturan asosiasi setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif A U B. Nilai *confidence* dari aturan A U B diperoleh dengan rumus berikut:

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

Rumus 2.2 Pembentukan Aturan Asosiasi

2.1.2.2. Pembelian dan Penjualan

Pembelian dan penjualan tidak bisa dipisahkan karena proses tersebut saling terhubung secara dependen. Pembelian adalah syarat untuk terjadinya proses penjualan, sehingga perusahaan akan mendapatkan keuntungan. Semakin banyak proses jual beli, semakin proporsional keuntungan yang didapat. Untuk membuat penjualan dilakukan untuk memastikan orang atau pihak lain memiliki keinginan untuk membeli produk yang ditawarkan (Lasmana Putra et al., 2019). Adapun hal-hal dianggap sebagai faktor berpengaruh pada penjualan sebagai berikut:

1. Kondisi Pasar (*market*)

Tempat jual beli antara penjual dan pembeli. Faktor kondisi pasar yaitu jenis kelompok pasar pembeli, segmentasi pasar, daya beli, frekuensi pembelian dan kebutuhannya agar penjual dapat menghasilkan pendapatan yang maksimal. Kondisi dan Kemampuan Penjual.

2. Modal

Ketika diawal proses usaha, penjual membutuhkan modal untuk dapat menunjukkan kepada pembeli bahwa produknya dijual melalui fasilitas komersial, seperti biaya pengiriman, biaya promosi barang, dan biaya lainnya.

3. Kondisi Organisasi Perusahaan

Di edukits yang termasuk usaha bisnis besar, penjualan akan ditangani oleh departemen pemasaran atau spesialis penjualan perusahaan. Berbeda dengan usaha kecil atau usaha kecil dimana proses penjualannya ditangani oleh satu manajemen atau pihak yang juga menangani banyak pekerjaan di dalam perusahaan atau bisnis tersebut.

2.1.2.3. Perangkat Lunak Tanagra

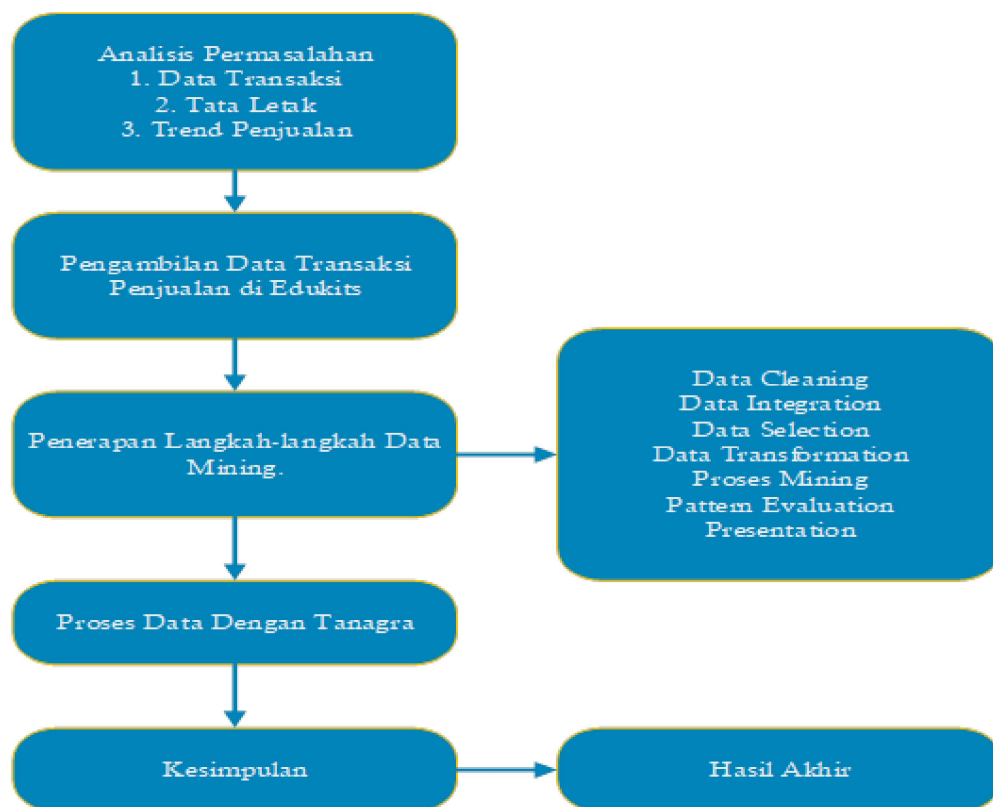
Software ini merupakan sebuah program dengan model *open-source* atau dengan sumber terbuka di mana siapa pun yang menggunakannya dapat langsung mengakses kode sumber program dan menambahkan algoritme sebagaimana diizinkan oleh perangkat lunak. Tujuan tanagra secara khusus (Zahara & Harman, 2021):

1. Tanagra digunakan untuk mencapai keadaan dimana peneliti bisa dimudahkan salah satu cara dengan software data mining.
2. Software ini juga untuk bertujuan untuk memberikan rekomendasi bagi para peneliti arsitektur agar mereka dapat menambahkan metode penambangan data ke data mereka dan membandingkan kinerjanya.
3. Kemudian guna dari perangkat ini yaitu untuk peneliti/pengembang pemula, yang pada gilirannya mempopulerkan metode yang digunakan untuk menghasilkan perangkat lunak sejenis. Mereka juga bisa mendapatkan keuntungan dari akses gratis ke kode sumber, yang memungkinkan mereka untuk melihat bagaimana perangkat lunak dibuat.

Proses pemilihan asosiasi untuk itu dilakukan pengurutan atau juga perhitungan berdasarkan skor dari *Support* juga nilai skor *Confidence* sehingga muncul rules atau Aturan dengan nilai terbesar untuk digunakan sebagai hasil atau rules asosiasi.

2.2. Kerangka Pemikiran

Penelitian ini akan dilaksanakan mengikuti prosedur analisis data seperti yang sudah ada pada proses atau langkah data minig. Namun peneliti ingin menuliskan bagaimana kerangka berpikir dari penelitian ini sehingga bisa mempermudah dalam membangun struktur penelitian dengan baik. Berikut adalah kerangka penelitian yang dibuat oleh peneliti



Gambar 2.1 Kerangka Pemikiran
(Sumber: Peneliti, 2023)

2.3. Penelitian Terdahulu

Melihat masalah dan judul penelitian yang akan diteliti, maka diperlukan adanya pemaparan tentang penelitian terdahulu guna mengungkapkan fenomena yang sama dalam sudut pandang yang berbeda sehingga diharapkan dapat memperkaya pengetahuan. Secara ringkas, penelitian terdahulu yang mendukung penelitian yang akan dilakukan dapat dilihat di bawah ini:

Tabel 2.1 Referensi peneliti dari Penelitian Terdahulu

PT.	Nama Peneliti	Judul Penelitian	Hasil
1	Rahmi, Mikola (2021)	menentukan pola pembelian pada customer (Rahmi & Mikola, 2020)	Dengan algoritma apriori, sangat efektif dalam menemukan item sembako yang paling sering terjual pada Join Mart, yaitu Minyak Garam dan penyedap rasa yang menghasilkan nilai <i>support</i> 20,00% dan <i>confidence</i> 100,00%.
2	Erlin Elisa (2018)	Market Base Analysis Pada Mini Market Ayu Dengan	Hasil dari penelitian ini menemukan nilai <i>support</i> dan <i>confidence</i> tertinggi adalah Minyak dan

PT.	Nama Peneliti	Judul Penelitian	Hasil
		Algoritma Apriori (Elisa, 2018)	Susu dengan nilai <i>support</i> 42,85% dan <i>confidence</i> 85,71%
3	Lasut, Rimbawan (2020)	Penerapan Algoritma Apriori (Lasut & Rimbawan Oprasto, 2020)	penjualan dengan menggunakan strategi bundling merupakan salah satu strategi yang tepat untuk meningkatkan penjualan di toko Violet Vape Store dibanding penjualan yang tidak menggunakan strategi bundling.
4	Suprianto Panjaitan., et al (2019)	Implementation of Apriori Algorithm for Analysis of Consumer Purchase Patterns (Panjaitan et al., 2019)	Dari hasil pengujian dengan memasukan nilai support dan drinking confidence yang berbeda yaitu adalah, jika minimum support adalah 15% dan tingkat kepercayaan adalah 50%,

PT.	Nama Peneliti	Judul Penelitian	Hasil
		(Panjaitan et al., 2019)	aturan 87 aturan akan menjadi dihasilkan
5	Tio Prasetya et al., (2021)	Analisis data transaksi terhadap pola pembelian konsumen menggunakan metode algoritma apriori (Prasetya et al., 2021)	Hasil dari algoritma Apriori terdapat 6 aturan asosiasi, nilai tertinggi confidence 0,64 dengan aturan yang didapat jika membeli item 165SDB maka membeli KARAKTER
6	Analisis Pembelian Produk...(2022)	Analisis pembelian ... (Nainggolan & Elisa, 2022)	Hasil final rules dari nilai support 27,86 % dan nilai confidence 68,00 % yaitu barang yang sering dibeli secara bersamaan yaitu elektronik untuk ruang tamu dan furniture ruang tamu.
7	Iska Heriyati Sigalingging, Rika Harman (2020)	Analisis Pola Pembelian Produk Menggunakan Algoritma Apriori	Penerapan algoritma apriori pada toko Al Barokah Mart dengan teknik data mining sangat efisien. Dengan

PT.	Nama Peneliti	Judul Penelitian	Hasil
		(Sigalingging & Harman, 2020)	support dan confidence tertinggi adalah Telur dan indomie dengan nilai support 46,66% dan nilai confidence 87,50%
8	Gaerry Setiawan, Dicky Surya Dwi Putra, Hartana Wijaya	Aplikasi Data Mining Berbasis Web Menggunakan Algoritma Apriori Untuk Analisa Pola Pembelian Barang Pada PT Menara Bahagia Bersama (Setiawan, 2012)	Aplikasi berbasis web yang dihasilkan melakukan analisa pola pembelian dari data transaksi yang dimasukkan, dengan cara menentukan range tanggal pada data yang ingin dianalisa, serta memasukkan nilai minimum support dan minimum confidence yang diinginkan
9	Ariefana Ria Riszky, Mujiono Sadikin (2019)	Data Mining Menggunakan Algoritma Apriori untuk Rekomendasi	Algoritma apriori dapat menghasilkan aturan asosiasi sebagai acuan dalam promosi produk dan

PT.	Nama Peneliti	Judul Penelitian	Hasil
		Produk bagi Pelanggan (Riszky & Sadikin, 2019)	pendukung keputusan dalam memberikan rekomendasi produk kepada konsumen berdasarkan nilai support dan confidence minimum yang telah ditetapkan.
10	Eni Heni Hermaliani et al., (2020)	Pola Penjualan & Persediaan Buah-buahan Menggunakan Algoritma Apriori (Hermaliani et al., 2020)	model perhitungan analisis pola frekuensi tinggi dan pembentukan aturan asosiasi dan pola kombinasi item set yang dihasilkan dari penjualan produk buah dari retail dengan dukungan dan kepercayaan tertinggi dari produk buah yang paling banyak terjual. Oleh karena itu algoritma apriori dapat membantu mengembangkan strategi penjualan

PT.	Nama Peneliti	Judul Penelitian	Hasil
11	Yusuf Kurnia, Yohanes shariato, Ceng Giap, Aditiya Hermawan, Riki	Study of application of data mining market basket analysis for knowing sales pattern (association of items) at the O! Fish restaurant using apriori algorithm (Kurnia et al., 2019)	Menganalisis pola pembelian (aturan asosiasi item) oleh konsumen di mana pola pembelian tersebut dapat digunakan sebagai rekomendasi dalam menentukan strategi pengembangan promosi untuk restoran
12	Primasatria Edastama, Ankur Singh Bist, Ari Prambudi	Implementation Of Data Mining On Glasses Sales Using The Apriori Algorithm (Edastama et al., 2021)	Penulis menggunakan algoritma Apriori untuk memperoleh informasi tentang barang-barang paling populer dan diminati dalam transaksi database di perusahaan optik nasional, sehingga hasilnya dapat digunakan untuk meningkatkan pertumbuhan

PT.	Nama Peneliti	Judul Penelitian	Hasil
			penjualan kaca mata dan pemasaran.
13	M. Kavitha, Dr. S. Subbaiah	Association Rule Mining using Apriori Algorithm for Extracting Product Sales Patterns in Groceries (Kavitha & Subbaiah, 2020)	Aturan asosiasi sangat berguna untuk menganalisis dataset yang dikumpulkan di supermarket, manajer dapat mengetahui produk apa yang sering dibeli bersamaan oleh pelanggan. Manajemen terbantu dalam pengambilan keputusan dan meningkatkan penjualan, memberikan penawaran kombinasi dan berbagai tawaran lainnya berdasarkan informasi yang didapatkan dari analisis aturan asosiasi.

PT.	Nama Peneliti	Judul Penelitian	Hasil
14	M.Hamdani Santoso	Application of Association Rule Method Using Apriori Algorithm to Find Sales Patterns Case Study of Indomaret Tanjung Anom (Santoso, 2021)	Hasil dari pengujian menggunakan metode algoritma Apriori menunjukkan adanya aturan asosiasi bahwa pelanggan sering membeli pasta gigi dan deterjen dengan memenuhi nilai kepercayaan minimum.
15	Risnamawati, Hasugian	Determination of Data Mining Application Design With Apriori Algorithm(Ndruru & Hasugian, 2020)	Data mining digunakan untuk menentukan urutan item makanan di Restoran Fountain beserta jumlah yang diperlukan.

(Sumber: Peneliti, 2023)