

## **BAB II**

### **LANDASAN TEORI**

#### **2.1. Pengertian Analisis**

Analisis adalah mengamati aktivitas objek dengan cara mendeskripsikan komposisi objek dan menyusun kembali komponen-komponennya untuk dikaji atau dipelajari secara detail. Kata analisis berasal dari bahasa Yunani Kuno ἀνάλυσις (analysis, "memecahkan" atau "menguraikan" dari ana- "naik, menyeluruh" dan lysis "melonggarkan").

Dalam bidang matematika, logika, analisis adalah proses pemecahan suatu masalah kompleks menjadi bagian-bagian kecil sehingga bisa lebih mudah dipahami. Dalam bidang kimia, analisis adalah penguraian suatu zat menjadi zat-zat yang lebih sederhana yang menjadi unsur-unsur pembentuknya.

Dalam linguistik, analisis adalah kajian yang dilaksanakan terhadap sebuah bahasa guna meneliti struktur bahasa tersebut secara mendalam. Sedangkan pada kegiatan laboratorium, kata analisis dapat juga berarti kegiatan yang dilakukan di laboratorium untuk memeriksa kandungan suatu zat dalam cuplikan.v

Bentuk baku dari istilah ini adalah analisis, yaitu serapan dari bahasa Inggris analysis. Akhiran -ysis bila diserap ke dalam bahasa Indonesia menjadi -isis. Bentuk analisis timbul karena pengaruh bahasa Belanda.

#### **2.2. Data Penjualan**

Data penjualan merupakan suatu proses yang bertujuan untuk menggali informasi dari data penjualan yang ada, dengan harapan dapat memberikan

wawasan yang berguna dalam pengambilan keputusan bisnis. Dalam konteks bisnis pupuk seperti yang dilakukan oleh Jaya Tani, data penjualan dapat membantu dalam memahami pola pembelian, preferensi konsumen, serta hubungan antara produk yang dibeli secara bersamaan. Dengan adanya pemahaman ini, perusahaan dapat merancang strategi pemasaran yang lebih efektif, meningkatkan stok barang yang tepat, serta memperbaiki layanan kepada pelanggan.

data penjualan juga dapat digunakan untuk meramalkan tren penjualan di masa depan. Salah satu pendekatan yang sering digunakan dalam data penjualan adalah dengan menggunakan metode *Data Mining*, yang dapat mengekstraksi pola dan informasi tersembunyi dalam data yang sangat besar dan kompleks. Salah satu teknik *Data Mining* yang populer adalah *Apriori Algoritma*.

### **2.2.1. Tujuan Data Penjualan**

Tujuan utama dari data penjualan meliputi:

1. Identifikasi Tren: Menemukan pola penjualan dari waktu ke waktu.
2. Segmentasi Pelanggan: Mengelompokkan pelanggan berdasarkan perilaku pembelian.
3. Optimasi Stok: Menentukan kapan dan berapa banyak produk yang harus disimpan.
4. Pengambilan Keputusan: Membantu manajemen dalam merencanakan strategi pemasaran dan penjualan.

### **2.3. Metode Apriori**

*Algoritma Apriori* adalah *Algoritma* yang digunakan untuk menghitung aturan asosiasi antar objek. Aturan asosiasi menjelaskan bagaimana dua atau lebih objek terkait satu sama lain. Dengan kata lain, *Algoritma apriori* adalah *Algoritma* berbasis aturan asosiasi yang menganalisis apakah orang yang membeli produk A juga membeli produk B. *Algoritma* ini dikemukakan oleh ilmuwan R. Agrawal dan Srikant pada tahun 1994. *Algoritma* ini kebanyakan digunakan untuk analisis keranjang belanja dan membantu menemukan produk yang dapat dibeli bersama.

*Apriori* juga dapat digunakan di bidang perawatan kesehatan untuk menemukan reaksi obat untuk pasien. Pada *Data Mining*, *Algoritma Apriori* banyak digunakan untuk menemukan data yang paling sering muncul dalam sebuah database. Item data transaksi pada database membentuk itemset. Itemset yang paling sering muncul dipilih oleh *Algoritma Apriori* agar dapat digunakan untuk menentukan aturan asosiasi yang menyoroti tren umum dalam database.

### **2.4. Knowledge Discovery in Database (KDD)**

Teknik *Data Mining* disebut dengan nama *Knowledge Discovery in Database* (KDD). *Knowledge Discovery in Database* (KDD) adalah istilah penerapan metode ilmiah pada *Data Mining* (Irawan, 2019) dalam penelitian (*Algoritma*, 2022). Dalam konteks ini, *Data Mining* merupakan langkah dalam proses *Knowledge Discovery in Database* (KDD). Prosedur *Knowledge Discovery in Database* (KDD) dipecah menjadi tahapan berikut:

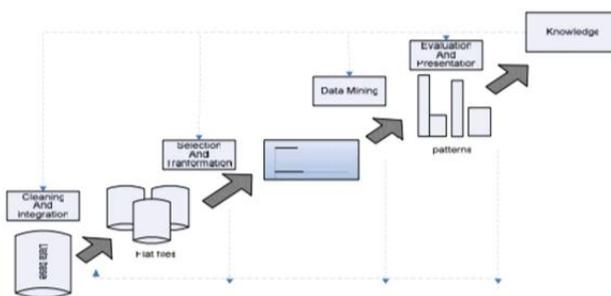
1. Informasi Seleksi (*selection*) Penting untuk mengumpulkan data (seleksi) dari sekumpulan data operasional sebelum memulai tahap *Knowledge Discovery*

*in Database* (KDD) dari ekstraksi informasi. Basis data operasional disimpan dalam file yang berbeda dari data pemilihan *Data Mining*.

2. Persiapan Data (Pembersihan/*Preprocessing*) Sebagai bagian dari prosedur pra-pemrosesan, duplikasi data dihilangkan, data yang tidak konsisten diperiksa, dan masalah data, seperti kesalahan ketik, diperbaiki. Proses “memperkaya” data yang sudah ada dengan lebih banyak data atau informasi, seperti data atau informasi eksternal, yang signifikan dan diperlukan untuk KDD juga dilakukan.
3. Transformasi Data yang sudah siap untuk proses *Data Mining* atau belum memiliki entitas tertentu berupa data yang valid diubah pada langkah ini.
4. Ekstraksi informasi dari data Dalam proses ini, teknik atau *Algoritma* pencarian informasi diterapkan.
5. Interpretasi / Penilaian
5. Fase terakhir ini melibatkan pembuatan output yang mudah dipahami dan didasarkan pada pola pengetahuan yang ditemukan selama proses penambangan data (A Ikhwan et al., 2020) dalam penelitian (*Algoritma*, 2022).

Informasi tentang penemuan pengetahuan dalam basis data (KDD), yang memiliki potensi untuk membantu penerapan penambangan data, dapat diperoleh dari sejumlah sumber terkini. Penambangan data, juga dikenal sebagai Penemuan Pengetahuan dalam Basis Data (KDD), adalah proses pengumpulan data historis dan menerapkannya untuk mencari pola, keteraturan, atau korelasi dalam kumpulan data yang sangat besar. Terdapat model atau mode yang digunakan untuk melakukan proses penggalian informasi dari data yang ada dalam

pengembangan teknologi *Data Mining*. Penambangan data, juga dikenal sebagai penemuan pengetahuan dalam basis data (KDD), mencoba memanfaatkan data basis data dengan mengelolanya untuk menghasilkan informasi baru yang relevan (Eko Prasetyo, 2014) dalam penelitian (*Algoritma*, 2022).



**Gambar 2. 1. Proses pada Knowledge Discovery in Database**

## 2.5. Data Mining

*Data Mining* menerapkan ilmu matematika, teknik statistic, dan kecerdasan buatan untuk menguraikan dan mengidentifikasi sebuah studi kasus dalam penelitian. Aplikasi *Data Mining* memanfaatkan data masukan berupa data penjualan, yang dapat menganalisis, meringkas dan mengekstrak data untuk menjadi sebuah informasi yang berguna akan tetapi tidak hanya mengandalkan data yang ada saja, perlu dilakukan analisis data sebelumnya untuk menggali potensipotensi yang ada dari data tersebut maka akan didapatkan pola dan informasi untuk pengambilan keputusan di waktu yang akan datang. *Algoritma apriori* adalah *Algoritma* paling terkenal untuk menemukan pola frekuensi tinggi. Pola frekuensi tinggi adalah pola-pola item di dalam suatu database yang memiliki frekuensi atau support di atas ambang batas tertentu yang disebut dengan istilah *minimum support*. *Algoritma apriori* dibagi menjadi beberapa tahap

yang disebut iterasi atau pass yaitu (Sianturi et all, 2018) dalam penelitian (Panjaitan, 2021). Berdasarkan penelitian terdahulu yang berkaitan dengan *Algoritma Apriori* (Mariani 2019) dalam penelitian (Panjaitan, 2021) yaitu “Implementasi *Algoritma Apriori* Pada Persediaan Barang Minuman Larutan Cap Kaki Tiga” hasil dari penelitian tersebut adalah Proses untuk menghasilkan kombinasi item menggunakan *Algoritma apriori* dilakukan dengan menetapkan nilai frequent itemset melalui kombinasi 1 itemset, 2 itemset, 3itemset dan 4 itemset. Untuk pembentukan asosiasi rule dilakukan dengan cara memberi nilai support 60% dan confidence 70%. Menurut Junaidi, (2019) dalam penelitian (Panjaitan, 2021) membahas tentang Implementasi *Algoritma Apriori* dan FP-Growth Untuk Menentukan Persediaan Barang, dalam penelitian ini support ditentukan menggunakan ambang batas 60% dan confidence 90%. *Apriori* telah berhasil menemukan sejumlah pola hubungan antar atribut dalam database retail. Penelitian yang dilakukan oleh Valerian & Hakim, (2018) dalam penelitian (Panjaitan, 2021) membahas tentang implementasi *Algoritma apriori* untuk prediksi stok peralatan tulis pada toko xyz. Hasil ini menunjukkan bahwa produk peralatan tulis yang sering dibeli oleh konsumen adalah Buku dan Pulpen dengan tingkat support 60% dan tingkat confident 80%. Berdasarkan penelitian-penelitian yang telah dilakukan sebelumnya akan dijadikan sebagai acuan atau pedoman dalam penelitian ini.

*Data Mining* mampu membantu perusahaan untuk mendapatkan pola dari data – data yang tersimpan di dalam basis data perusahaan. Pengetahuan yang diperoleh tersebut akan menjadi pedoman dalam mengambil tindakan – tindakan

bisnis sebagai upaya pemeliharaan dan peningkatan fungsi – fungsi berikut : deskripsi, estimasi, prediksi, klasifikasi, clustering, dan asosiasi (Nurjoko dan Kurniawan, 2016). *Data Mining* juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan selama ini tidak diketahuai secara manual dari sekumpulan data (Pramudiono, 2007. Dalam buku Kusriani, dkk, 2009). *Data Mining*, sering juga disebut sebagai Knowledge Discovery in Database (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, histori untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar (Santoso, 2007) dalam penelitian (Panjaitan, 2021)

Dengan mengekstraksi dan mengidentifikasi pola signifikan dari basis data, *Data Mining* adalah serangkaian prosedur yang bertujuan untuk mengungkap nilai tambahan berupa informasi yang sebelumnya tidak ditemukan dari basis data (Amrin Amrin, 2017). Penambahan penambangan data adalah pemeriksaan pengamatan basis data dengan tujuan menemukan asosiasi yang tidak terduga serta untuk meringkas data dalam bentuk baru atau mudah dipahami yang berguna bagi pemilik data. Kumpulan data dapat ditambang untuk informasi yang dapat dikumpulkan dan digunakan. Informasi yang dikumpulkan meliputi keterkaitan antar item data, pola tersembunyi dalam data, dan pengembangan model untuk peramalan data (Adinugroho & Sari, 2018).

Analisis otomatis kumpulan data besar atau rumit dengan tujuan mengidentifikasi tren atau kecenderungan signifikan yang biasanya diabaikan dalam kehidupan sehari-hari dikenal sebagai *Data Mining*. *Data Mining* studi tentang memeriksa kumpulan data untuk menemukan asosiasi yang tidak terduga

dan meringkas data dalam bentuk yang jelas bagi pemilik data (Ali Ikhwan et al., 2020). Memproses volume data yang sangat besar untuk mendapatkan pengetahuan atau informasi yang membutuhkan suatu metode dikenal sebagai *Data Mining*. *Data Mining* harus berperan dalam proses mengubah data mentah menjadi informasi, pola, dan keahlian yang berharga. *Data Mining* umumnya terdiri dari lima peran: perkiraan, prediksi, klasifikasi, pengelompokan, dan asosiasi (Suendri et al., 2020). Data merupakan gabungan fakta yang tersimpan, *Mining* merupakan suatu proses penambangan (Ali Ikhwan et al., 2021). *Data Mining* adalah proses mencari pola yang belum ditemukan dalam sekumpulan data yang disimpan 7 dalam database, gudang data, atau jenis media penyimpanan informasi lainnya. *Data Mining* berupaya mengungkap pengetahuan yang masih tersembunyi dalam kumpulan data. untuk menemukan, menyelidiki, atau menambang pengetahuan dari fakta atau informasi sendiri.(Panjaitan, 2021)

## **2.6. Association rule**

*Association rule* adalah teknik *Data Mining* yang berguna untuk menemukan suatu korelasi atau pola yang terpenting atau menarik dari sekumpulan data besar. *Association rules* merupakan salah satu metode yang bertujuan mencari pola yang sering muncul di antara banyak transaksi, dimana setiap transaksi terdiri dari beberapa item sehingga metode ini akan mendukung system rekomendasi melalui penemuan pola antar item dalam transaksi-transaksi yang terjadi. Menurut beberapa jurnal yang telah dipublikasi sebelumnya aplikasi *Data Mining* dapat digunakan untuk menampilkan informasi tingkat kelulusan. Informasi yang ditampilkan berupa nilai support dan confidence hubungan antara

tingkat kelulusan dengan data induk mahasiswa. Semakin tinggi nilai confidence dan support maka semakin kuat nilai hubungan antar atribut. Penerapan *Association rule Algoritma apriori* dalam aplikasi business analytic terhadap data lulusan dari tahun 2005 sampai 2013 di UNS, dapat digunakan untuk membantu mendapatkan informasi tentang kualitas kelulusan, serta untuk pengambilan keputusan dalam meningkatkan kualitas kelulusan. *Data Mining* dapat diimplementasikan dengan menggunakan database penjualan alat-alat kesehatan karena dapat menemukan kecenderungan pola kombinasi itemsets sehingga dapat dijadikan sebagai informasi yang sangat berharga dalam pengambilan keputusan untuk mempersiapkan stok jenis barang yang diperlukan kemudian. Penerapan *Algoritma apriori* pada Teknik *Data Mining* sangat efisien dan dapat mempercepat proses pembentukan kecenderungan pola kombinasi itemset. Berdasarkan kebutuhan dan review jurnal sebelumnya, *Data Mining* market basket analysis merupakan solusi yang dapat digunakan untuk diimplementasikan (Ulumuddin et al., 2018).

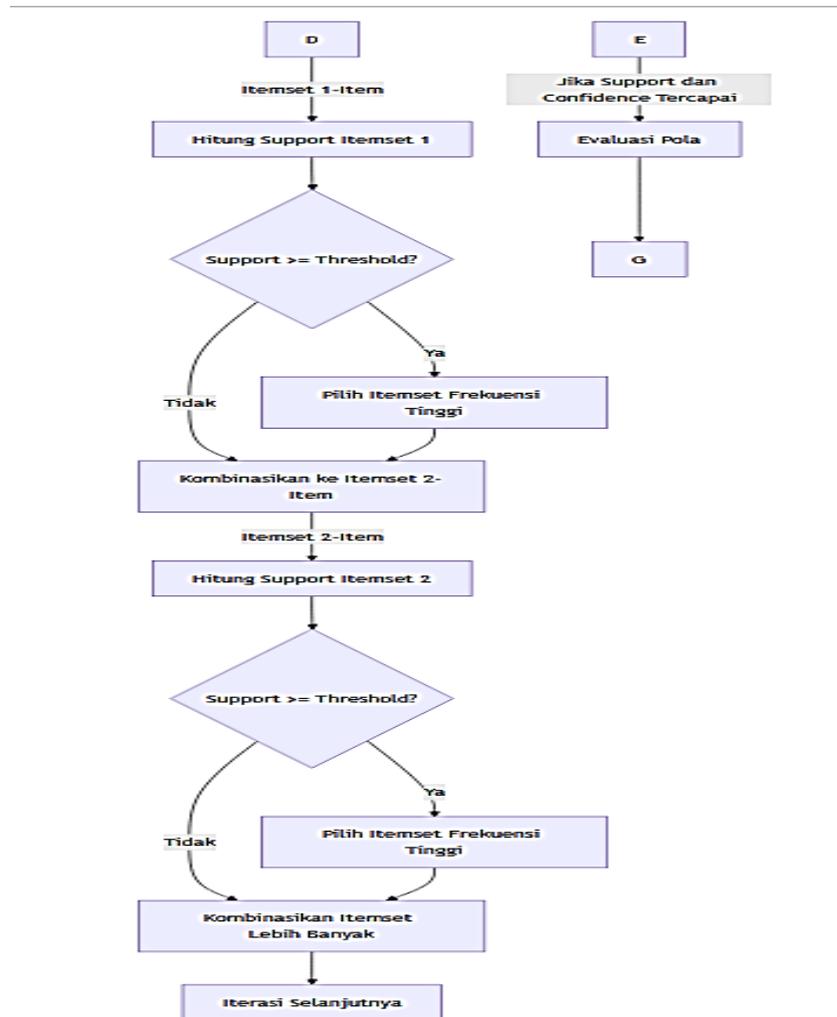
### **2.7. Algoritma Apriori**

*Algoritma Apriori* merupakan salah satu *Algoritma* yang paling banyak digunakan dalam *Data Mining* untuk mencari pola asosiasi dalam dataset yang besar. Pada *Algoritma Apriori*, dua parameter utama yang digunakan untuk menentukan relevansi suatu itemset adalah *minimum support* dan *minimum confidence*. Support mengacu pada proporsi transaksi yang harus mencakup itemset tertentu agar dianggap signifikan dalam dataset, sementara confidence mengukur seberapa sering aturan asosiasi yang dihasilkan benar (Rizaldi &

Adnan, 2021) [9]. Kedua parameter ini membantu dalam menyaring itemset dan aturan yang tidak relevan, sehingga hanya pola-pola yang kuat dan bermakna yang dipertimbangkan untuk analisis lebih lanjut.

Setelah parameter support dan confidence ditentukan, *Algoritma Apriori* mulai mencari frequent itemsets, yaitu kombinasi item yang sering muncul bersama dalam transaksi. Proses ini dilakukan secara iteratif, dimulai dengan itemset tunggal dan kemudian menggabungkan item-item tersebut untuk menemukan kombinasi itemset yang lebih besar. Pada setiap iterasi, *Algoritma* menghitung support untuk itemset yang dihasilkan dan menyaring itemset yang tidak memenuhi kriteria *minimum support* [10] (Nurajizah et al., 2023). Dengan demikian, *Algoritma Apriori* dapat mengidentifikasi aturan asosiasi yang menunjukkan hubungan antara produk-produk yang sering dibeli bersama, yang sangat berguna untuk strategi pemasaran dan pengelolaan inventaris (Ardianto & Fitriyah, 2019) (Irfiani, 2019).

### 2.7.1. Langkah Kerja *Algoritma Apriori*



Gambar 2. 2. Langkah Kerja *Algoritma Apriori*

### 2.7.2. Formula Matematika yang Digunakan

Dalam *Algoritma Apriori*, terdapat tiga konsep utama yang digunakan untuk mengukur kualitas dan kekuatan dari aturan asosiasi yang ditemukan, yaitu support, confidence, dan lift.

Support: Support mengukur seberapa sering itemset muncul dalam dataset.

Support untuk itemset A dihitung dengan rumus berikut:

$$Support(A) = \frac{Jumlah\ transaksi\ yang\ mengandung\ itemset\ A}{Jumlah\ total\ transaksi} \quad (1)$$

Support yang tinggi menunjukkan bahwa itemset tersebut sering muncul dalam data, yang mengindikasikan bahwa itemset tersebut penting dalam analisis.

Confidence: Confidence mengukur probabilitas bahwa jika item A diambil, maka item B juga akan diambil. Rumus untuk menghitung confidence adalah:

$$\text{Confidence}(A \rightarrow B) = \frac{\text{Support}(A \cup B)}{\text{Support}(A)} \quad (2)$$

Confidence yang tinggi menunjukkan bahwa ada kemungkinan besar bahwa ketika item A diambil, item B juga akan diambil. Ini menggambarkan hubungan yang kuat antara dua item dalam aturan asosiasi.

Lift: Lift mengukur kekuatan hubungan antara dua item yang diambil bersama dibandingkan dengan yang diharapkan jika mereka diambil secara independen. Lift dihitung dengan rumus berikut:

$$\text{Confidence}(A \rightarrow B) = \frac{\text{Support}(A \rightarrow B)}{\text{Support}(B)} \quad (3)$$

Lift yang lebih besar dari 1 menunjukkan bahwa hubungan antara A dan B lebih kuat dibandingkan dengan apa yang diharapkan jika keduanya muncul secara independen. Metrik ini sangat berguna untuk menilai seberapa signifikan asosiasi antara dua barang.

## **2.8. RapidMiner**

RapidMiner merupakan perangkat lunak yang bersifat terbuka (*open source*). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap *Data Mining*, text mining dan analisis prediksi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. RapidMiner memiliki kurang lebih 500 operator *Data Mining*, termasuk operator untuk input,

output, data preprocessing dan visualisasi. RapidMiner merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin *Data Mining* yang dapat diintegrasikan pada produknya sendiri. RapidMiner ditulis dengan menggunakan bahasa java sehingga dapat bekerja di semua sistem operasi.

a. RapidMiner memiliki beberapa sifat sebagai berikut:

1. Ditulis dengan bahasa pemrograman Java sehingga dapat dijalankan di berbagai sistem operasi.
2. Proses penemuan pengetahuan dimodelkan sebagai operator trees.
3. Representasi XML internal untuk memastikan format standar pertukaran data.
4. Bahasa scripting memungkinkan untuk eksperimen skala besar dan otomatisasi eksperimen.
5. Konsep multi-layer untuk menjamin tampilan data yang efisien dan menjamin penanganan data.
6. Memiliki GUI, command line mode, dan Java API yang dapat dipanggil dari program lain.

b. Beberapa Fitur dari RapidMiner, antara lain:

1. Banyaknya *Algoritma Data Mining*, seperti decision treee dan self-organization map.
2. Bentuk grafis yang canggih, seperti tumpang tindih diagram histogram, tree chart dan 3D Scatter plots.
3. Banyaknya variasi plugin, seperti text plugin untuk melakukan analisis teks.
4. Menyediakan prosedur *Data Mining* dan machine learning termasuk: ETL (extraction, transformation, loading), data preprocessing, visualisasi,

modelling dan evaluasi<sup>5</sup>. Proses *Data Mining* tersusun atas operator-operator yang nestable, dideskripsikan dengan XML, dan dibuat dengan GUI<sup>6</sup>. Mengintegrasikan proyek *Data Mining* Weka dan statistika R (Rasianto & Sutedi, 2023).



Gambar 2. 3. RapidMiner

## 2.9. Metode Penelitian

Pada penelitian ini metode penelitian akan menjelaskan bagaimana penelitian terdahulu dalam menggunakan metode yang digunakan pada penelitian ini.

### 2.9.1. Penelitian terdahulu

Dibawah ini adalah refrensi yang penulis gunakan mendukung dan sebagai landasan pembuatan proposal yaitu sebagai berikut:

Tabel 2. 1. Penelitian Terdahulu

Refrensi penelitian	1 (Arliana Nur Kadim, 2023)
Judul	PENERAPAN <i>ALGORITMA APRIORI</i> MENENTUKAN KORELASOI DATA PENJUALAN PUPUK (STUDI KASUS : PT. KURNIA ROTORINDO TANI)
Nama penulis	Ardianti

Tahun	2023
Hasil	Pada refrensi jurnal ini di jelaskan penelitian ini bertujuan untuk mengidentifikasi pola pembelian pupuk berdasarkan data transaksi penjualan. Dengan <i>Algoritma Apriori</i> , diharapkan dapat ditemukan korelasi antar produk pupuk yang sering dibeli bersamaan sehingga dapat membantu perusahaan dalam mengambil keputusan, seperti strategi pemasaran, penempatan produk, atau penyusunan promosi.
Refrensi penelitian	2 (Panjaitan, 2021)
Judul	IMPEMENTASI <i>ALGORITMA APRIORI</i> UNTUK PENENTUAN POLA PERSEDIAAN BARANG PADA UD.CHANDRA JAYA TANI
Nama penulis	Fitra Sari Penjaitan
Tahun	2021
Hasil	Pada refrensi jurnal ini dijelaskan mengenai tujuan Penelitian untuk iplementasi <i>Algoritma apriori</i> berhasil mengidentifikasi pola pembelian yang signifikan dan dapat di gunakan sebagai dasar pengeambilan keputusan dalam memamanajemen persediaan barang di UD. Chandra jaya tani, sebuah toko pertanian
Refrensi penelitian	3 (Nuryahya et al., 2024)
Nama Penulis	Muhamad Nuryahya Dan Andri Triyono
Judul	IMPLEMENTASI <i>ALGORITMA APRIORI</i> UNTUK MENCARI POLA TRANSAKSI PENJUALAN PADA TOKO PERTANIAN TOKO BIDSALTANI
Tahun	2024

Hasil	Pada refrensi jurnal ini di jelaskan mengenai tujuan penelitian untuk Implementasi <i>Algoritma Apriori</i> dalam menganalisis pola transaksi penjualan di Toko Pertanian BIDSALTANI berhasil mengidentifikasi berbagai asosiasi antar produk yang sering dibeli bersamaan. Dengan hasil ini, toko dapat mengoptimalkan strategi pemasaran dan manajemen stok dengan lebih efisien. Penggunaan <i>Algoritma Apriori</i> dapat menjadi alat yang sangat berguna dalam meningkatkan kinerja penjualan dan kepuasan pelanggan.
Refrensi penelitian	4
Judul	Penerapan Metode <i>Association rule</i> Mining (ARM) Untuk Memprediksi Rencana Penambahan Stok Pupuk Berdasarkan Kebiasaan Pelanggan Pada Kelompok Tani Di Desa Bumisari Kecamatan Natar
Nama Penulis	Rasianto Dan Sutedi
Tahun	2023
Hasil	Pada refrensi jurnal ini dijelaskan Penerapan metode <i>Association rule</i> Mining (ARM) dengan <i>Algoritma Apriori</i> pada data transaksi pembelian pupuk oleh petani di Desa Bumisari berhasil mengidentifikasi pola-pola pembelian pupuk yang bergantung pada jenis tanaman yang ditanam. Hasil analisis ini dapat digunakan untuk merencanakan penambahan stok pupuk yang lebih tepat sasaran, sehingga membantu petani mendapatkan pupuk yang dibutuhkan tepat waktu dan dengan jumlah yang sesuai.

## 2.10. Metodologi Penelitian

*Data Mining* (Penggalian data)

*Data Mining* (penggalian data) merupakan proses kesekian serta ikatan timbal balik guna memperoleh pola ataupun model baru yang sempurna,

bermanfaat, serta gampang dimengerti dalam basis data yang lumayan sangat besar [8].

#### *Association rule*

*Association rules* merupakan guna mengenali ketentuan asosiasi antara satu set elemen. Metodologi *Association rules* Informasi Mining dibagi jadi 2 sesi ialah [9]:

##### *Analisis Pola Frekuensi Tinggi (High Frequency Pattern Analysis)*

Analisis pola frekuensi tinggi ialah buat mencari gabungan antara itemset yang sudah terpenuhi ketentuan minimum dari nilai support tersebut. Nilai support suatu item didapatkan rumus selaku berikut:

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

Rumus support maksudnya nilai support dihitung dengan membagi jumlah yang mempunyai kandungan item A (satu item) dengan total seluruh transaksi. Sebaliknya nilai support dari 2 item didapatkan rumus selaku berikut:

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

Rumus support di atas berarti bahwa ketika nilai support dua itemset diperoleh dengan membagi jumlah transaksi yang berisi item A serta item B (item awal ditambah yang lain) dengan besaran total seluruh transaksi.

Pembuatan Aturan Asosiatif (*Formation of Associative Rules*) Sesudah menemukan acuan frekuensi tinggi, cari aturan asosiatif yang akan memenuhi persyaratan confidence minimum dengan menghitung confidence ketentuan asosiatif  $A \rightarrow B$ . Nilai confidence dari ketentuan  $A \rightarrow B$  didapatkan rumus selaku berikut:

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

Rumus di atas berarti bahwa ketika mencari nilai confidence pada aturan asosiasi  $A \rightarrow B$ , jumlah transaksi yang mengandung item A dan B dapat dibagi dengan jumlah transaksi yang mengandung A.

*Apriori Algorithm (Algoritma Apriori)* Algoritma apriori/ analisis asosiasi ialah Algoritma pengutipan informasi yang memakai ketentuan asosiatif yang memenuhi batasan support beserta confidence untuk membangun hubungan untuk satu objek yang sesuai dengan kriteria persyaratan minimum untuk nilai support dan confidence dengan membangun aturan asosiasi [10]. Algoritma Apriori dipecah sebagian sesi yang diujarkan literasi, yaitu sebagai berikut.

1. Membentuk kandidat itemset.

Kandidat k-itemset tercipta dari campuran(k- 1) itemset yang diperoleh pada itemset lebih dahulu. Salah satu watak dari Algoritma apriori merupakan guna memotong kandidat k-itemset yang himpunan bagiannya sudah memiliki k-1 item serta tidak tercantum dalam pola frekuensi besar dengan panjang k-1. 2) Hitung support dari tiap calon k-itemset Hitung k-itemset terpanjang dengan memindai seluruh database sebanyak mungkin, dan hitung jumlah transaksi yang muat seluruh item dalam kandidat k- itemset. 3) Tetapkan pola frekuensi besar

Penetapan pola frekuensi besar yang mengandung k-itemsets ditentukan oleh kandidat k-itemsets dengan dukungan lebih besar dari dukungan minimum. 4. Tidak mendapatkan pola frekuensi besar Bila tidak didapatkan pola frekuensi besar, maka k ditambahkan satu dan mengulang ke nomor 1. 4) Bahasa R R adalah bahasa pemrograman open source untuk menangani komputasi dan manipulasi data untuk menghasilkan statistik yang sangat membantu untuk penelitian dan industri [12]. R adalah sistem analisis statistik yang relatif lengkap, yang merupakan hasil penelitian kooperatif oleh ahli statistik dari seluruh dunia.

4) Tahapan Metode Analisis Data Berikut ini adalah tahapan prosedur analisis data memakai CRISP- DM (*Cross- Industry Standard Process for Data Mining*) ada 6 tahapan :

a. Pemahaman Bisnis (*Business Understanding*) PT. Pusri Palembang Jateng 1 dapat memenuhi kebutuhan konsumen dengan menyediakan berbagai pupuk non subsidi. PT. Pusri Palembang Jateng 1 mencari kombinasi pupuk dari data transaksi penjualan pupuk non subsidi untuk menemukan kombinasi pupuk yang paling umum sehingga pemilik perusahaan dapat menggunakannya sebagai rekomendasi untuk memberikan jenis pupuk prioritas berdasarkan kombinasi pupuk yang dijual dengan Gabungan dengan persediaan pengendalian pupuk non subsidi..

b. Pemahaman Data (*Data Understanding*) Tahap pengumpulan data transaksi dari PT. Pusri Palembang Jateng 1 dengan mengambil sampel data bulan Januari – Juli 2021. Data yang digunakan adalah dalam bentuk Microsoft Excel dengan jumlah data pada bulan Januari – Juli 2021 sebesar 744 record. Data

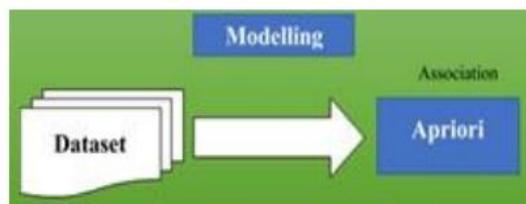
tersebut berisikan tabel penjualan yang terdiri dari 10 field. Field data transaksi penjualan PT. Pusri Palembang Jateng 1 bisa dilihat di tabel 1.

Tabel 2. 2. Field Data Transaksi Penjualan PT. Pusri Palembang Jateng 1

No	Nama Field
1	No Urut
2	Tanggal
3	No Skpp
4	No Po
5	Tgl Po
6	Distributor
7	Produk
8	Kemasan
9	Fot
10	Kuantitas

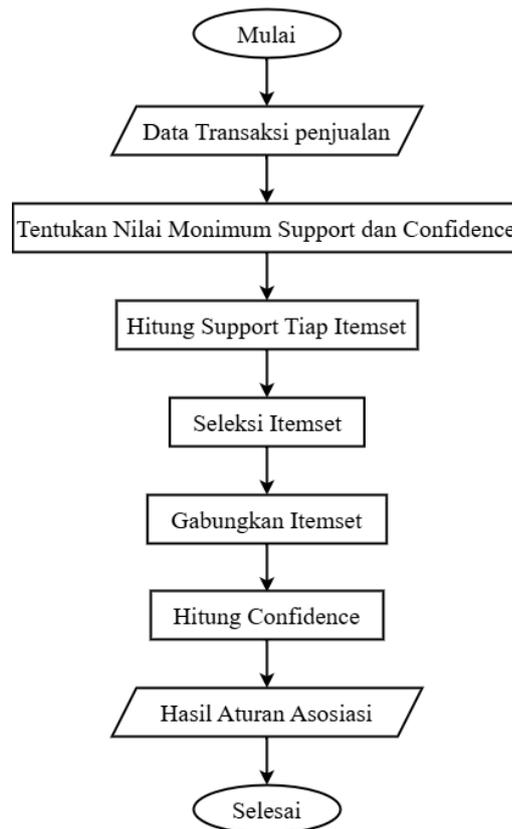
c. *Persiapan Data (Data Preparation)* Pada sesi ini dicoba persiapan data dengan memilih field yang akan dipakai kemudian ditransformasikan.

d. *Modelling Phase* Ditahap pemodelan penelitian menggunakan metode *Data Mining* dengan prosedur asosiasi *Algoritma apriori* yang dibantu dengan Rstudio. Proses ini menghasilkan pola pembelian pupuk dan memberikan informasi kombinasi pembelian pupuk yang saling berhubungan.



Gambar 2. 4. Pemodelan

*Algoritma Apriori* Flowchart proses *Algoritma Apriori* bisa dilihat di gambar 3 flowchat *Algoritma apriori*.



Gambar 2. 5. Flowchart Algoritma Apriori [13]

Berikut ini adalah penjelasan dari tahapan flowchart *Algoritma apriori* :

Tabel 2. 3. Penjelasan Tahapan Algoritma Apriori

No	Tahapan	Keterangan
1	Pengolahan Data	Mengolah data agar siap digunakan analisa data memakai <i>Algoritma apriori</i> .
2	Pencarian Calon Kombinasi (Pembentukan Itemset)	Mncari calon kombinasi dengan melakukan pengombinasian item dalam transaksi dengan kombinasi 1 itemset,2 itemset dst.

3	Kombinasi yang Terpilih	Hasil dari kombinasi yang terpilih berdasarkan <i>minimum support</i> dan <i>confidence</i> berupa rule
4	Selesai	maka proses selesai

e. Fase Evaluasi (*Evaluation Phase*) Tahap evaluasi melibatkan penentuan apakah hasil dari analisis *Algoritma apriori* cocok dengan kebutuhan toko. f. Fase Penyebaran (*Deployment Phase*) Hasil penelitian akan disampaikan dalam bentuk laporan kepada PT. Pusri Palembang Jawa Tengah 1 yang akan digunakan sebagai rekomendasi pupuk non subsidi untuk meningkatkan penjualan di PT. Pusri Palembang, Jawa Tengah1 [14].