

BAB II

TINJAUAN PUSTAKA

2.1 Landasan Teori

Landasan teori merupakan suatu teori yang di kutip dari buku-buku atau jurnal dan di jadikan sebagai referensi bagi peneliti untuk memperkuat hasil penelitian sehingga hasil yang akan di peroleh oleh penelith lebih akurat atau valid.

2.1.1 Pengertian Penjualan

Kegiatan penjualan merupakan bagian dari operasi suatu perusahaan atau pengusaha yang bertujuan untuk mempertahankan dan mengembangkan bisnis dengan mencapai keuntungan atau laba yang diinginkan. Penjualan adalah proses dari menetapkan harga jual hingga produk disalurkan kepada konsumen.

Penjualan adalah bagian penting yang melengkapi atau melengkapi aktivitas pembelian, dimana keduanya diperlukan agar transaksi dapat terjadi. Kegiatan pembelian dan penjualan adalah dua hal yang saling terhubung dan berkaitan untuk mempermudah proses transfer transaksi. Dengan demikian, kegiatan penjualan, seperti halnya kegiatan pembelian, melibatkan serangkaian langkah termasuk permintaan, mencari pembeli, negosiasi harga, syarat pembayaran, dan prosedur lainnya, yang memungkinkan pelaksanaan rencana penjualan yang telah ditetapkan (Djamaludin et al., 2017).

2.1.2 Pengertian Data Transaksi

Data transaksi merujuk pada data yang mencatat peristiwa-peristiwa yang telah terjadi dalam suatu usaha atau perusahaan, yang mencakup data penjualan, pengiriman, hutang, dan peristiwa lainnya. Data transaksi biasanya dapat dikategorikan dalam tiga kelompok berdasarkan kata kerja yang digunakan, yaitu keuangan untuk pesanan dan pembayaran, tenaga kerja untuk jadwal dan catatan kerja, serta logistik untuk pengiriman.

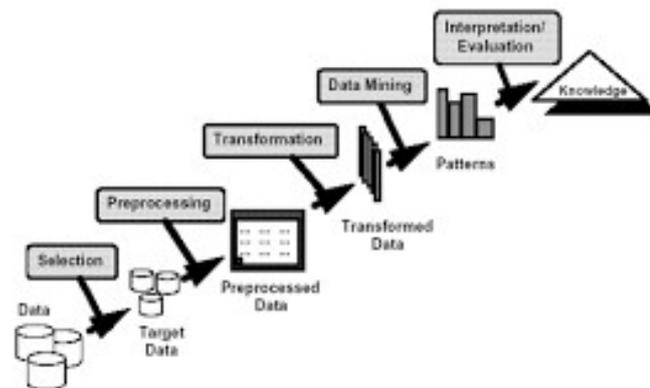
Definisi transaksi secara umum mencakup setiap aktivitas yang mempengaruhi aset atau keuangan organisasi atau individu. Beberapa contoh aktivitas transaksional meliputi penjualan, pembelian, pembayaran gaji, dan sebagainya. Dalam data transaksi, terdapat manajemen transaksi yang bertugas mencatat perubahan keuangan dengan cermat menggunakan metode tertentu. Data transaksi penjualan memiliki nilai penting dalam pengambilan keputusan bisnis (Djamaludin et al., 2017).

2.1.3 KDD (Knowledge Discovery in Database)

Metode yang dapat digunakan untuk melakukan data mining ialah “*Knowledge Discovery in Database Process*” (KDD). Di bawah ini disajikan gambaran dan penjelasan mendetail tentang proses KDD (Elyas & Prayoga, 2020):

1. *Data Cleansing* merupakan tahap pengolahan data yang berfokus pada pemilihan data yang dianggap relevan dan dapat digunakan dalam proses lebih lanjut.

2. *Data Integration* adalah proses penggabungan data yang dianggap memiliki duplikasi atau kesamaan sehingga dapat dikombinasikan menjadi satu kesatuan.
3. *Selection* adalah tahap seleksi atau penyortiran data yang relevan dan krusial untuk dipergunakan dalam analisis.
4. *Data Transformation* merupakan proses dimana data yang telah difilter mengalami perubahan bentuk sesuai dengan standar *data mining* yang akan dimanfaatkan.
5. *Data Mining* merupakan suatu tahapan di mana berbagai teknik diterapkan untuk mengekstrak pola potensial dari data sehingga mengeluarkan *output* berupa informasi yang memiliki nilai dan manfaat bagi pengguna.
6. *Pattern Evolution* adalah proses di mana pola-pola yang telah diidentifikasi dianalisis dan dievaluasi berdasarkan ukuran atau metrik yang telah ditentukan sebelumnya.
7. *Knowledge Presentation* adalah tahap terakhir dalam proses KDD, data yang sudah melewati pemrosesan direpresentasikan dengan tujuan agar lebih mudah dimengerti oleh pengguna, dan diekspektasikan dapat memberikan wawasan dan tindakan berlandaskan hasil analisis yang sudah dilaksanakan.



Gambar 2. 1 Tahapan KDD
Sumber: (Elyas & Prayoga, 2020)

2.1.4 Pengertian Data Mining

Data mining ialah sebuah proses pemahaman dan identifikasi informasi berharga dari basis data yang sangat besar memanfaatkan metode seperti statistik, matematika, kecerdasan buatan, dan pembelajaran mesin. Dalam “*Knowledge Discovery in Databases*” (KDD), terdapat berbagai pendekatan yang berbeda untuk mencari informasi atau pengetahuan. Pendekatan kuantitatif meliputi pencarian probabilistik seperti logika induktif, pencarian pola, dan analisis pohon keputusan. Selain itu, terdapat pendekatan analisis kecenderungan, deviasi, algoritma genetik, jaringan saraf tiruan, serta pendekatan kombinasi dari dua atau lebih metode sebelumnya.

Menurut (Fajrin & Handoko, 2018) *data mining* ialah eksplorasi sistematis dari kumpulan data yang besar yang telah disimpan sebelumnya, dengan memanfaatkan metode pengenalan pola statistik dan matematika. Tujuan utamanya adalah untuk menemukan hubungan, pola, dan kecenderungan yang bermakna dalam data tersebut.

2.1.5 Operasi Dasar *Data Mining*

Pada dasarnya, dalam *data mining* terdapat dua kategori sistem operasi yang dapat dikelompokkan seperti berikut:

1. Metode deskriptif bertujuan untuk mengidentifikasi pola, hubungan, atau anomali dalam data dengan cara yang lebih mudah dipahami oleh manusia, contohnya termasuk teknik *clustering* dan *association rule*.
2. Metode prediktif bertujuan untuk melakukan estimasi nilai dari suatu variabel lainnya, seperti klasifikasi dan regresi.

2.1.6 Fungsi *Data Mining*

Jika dilihat dari fungsinya, data mining dapat dikelompokkan menjadi beberapa bagian dalam menjalankan tugas-tugasnya. *Data mining* melibatkan berbagai tugas yang dapat dikelompokkan berdasarkan fungsinya. Pertama, terdapat klasifikasi, yang merupakan proses penggeneralisasian ketentuan yang sudah didapat untuk diterapkan di data baru, seperti mengklasifikasikan jenis penyakit berdasarkan gejala. Kedua, klusterisasi, yang bertujuan mengkalasifikasikan data yang belum diketahui label kelasnya menjadi kelompok-kelompok tertentu berdasarkan kemiripannya. Kemudian, terdapat regresi yang fokus pada menemukan fungsi untuk memodelkan data dengan prediksi seminimal mungkin. Selain itu, ada deteksi anomali yang berperan dalam mengidentifikasi data yang tidak umum, seperti outlier atau perubahan yang memerlukan penyelidikan lebih lanjut. Pembelajaran asosiasi atau pemodelan kebergantungan adalah proses mencari hubungan antara variabel dalam data. Terakhir,

perangkuman memberikan representasi data yang lebih sederhana melalui visualisasi dan pembuatan laporan untuk memahami informasi dengan lebih baik.

2.1.7 Algoritma Data mining

Beberapa algoritma ranking teratas diumumkan pada konferensi internasional terbesar tentang data mining, yaitu IEEE ISDM tahun 2006, seperti berikut ini:

2.1.7.1 Algoritma C4.5

C4.5 adalah teknik klasifikasi prediksi yang mampu mengubah dataset yang besar menjadi pohon keputusan yang mengungkapkan aturan-aturan dengan cara yang mudah dipahami. Aturan-aturan tersebut dapat diwujudkan dalam bentuk basis data yang mirip dengan SQL untuk mencari catatan pada kategori tertentu. Selain itu, C4.5 juga berguna dalam melakukan eksplorasi data dan menemukan hubungan tersembunyi antara berbagai variabel calon input dengan variabel target (Pambudi & Setiawan, 2018).

Dalam pemilihan atribut untuk menjadi akar, dilakukan berdasarkan nilai gain tertinggi dari berbagai atribut yang ada. Untuk menghitung gain tersebut, digunakan rumus sebagai berikut:

$Gain(S, A) =$

$$Entropy(S) \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad \text{Rumus 2. 1 Menghitung gain}$$

Keterangan:

S: Kumpulan kasus atau dataset.

A: Variabel atau ciri.

n: Banyaknya pembagian atau partisi dari atribut A.

$|S_i|$: Banyaknya data atau kasus dalam partisi ke- i .

$|S|$: Jumlah data atau kasus dalam himpunan S .

Sebelum mengkomputasi gain dengan mencari entropy, digunakan konsep entropy untuk menilai seberapa banyak informasi yang terkandung dalam sebuah atribut ketika menghasilkan atribut lainnya. Rumus dasar untuk menghitung entropy tersebut adalah sebagai berikut:

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i \quad \text{Rumus 2. 2 Entropy}$$

Keterangan:

S : Kumpulan kasus atau dataset.

N : Banyaknya pembagian atau partisi dalam himpunan S .

P_i : Rasio antara jumlah elemen dalam setiap partisi S_i terhadap jumlah keseluruhan elemen dalam himpunan S .

2.1.7.2 Algoritma K-means

Algoritma K-means adalah teknik klasterisasi yang bertujuan untuk mempartisi data yang ada menjadi satu atau lebih kelompok (klaster) berdasarkan kesamaan mereka. Algoritma ini mengelompokkan data dengan karakteristik serupa dalam satu klaster, sementara data dengan karakteristik yang berbeda ditempatkan dalam klaster terpisah (Studi et al., n.d.).

Proses algoritma clustering menggunakan metode K-Means melibatkan beberapa tahapan yang perlu diikuti (Adiya & Desnelita, 2019). Tahapan pertama adalah memilih jumlah cluster k yang diinginkan. Kemudian, tahapan kedua berfokus pada inisialisasi nilai awal untuk setiap cluster yang dipilih. Pada tahapan

ketiga, seluruh data atau objek akan dialokasikan ke cluster yang paling dekat berdasarkan jarak atau kesamaan karakteristiknya. Dengan demikian, proses K-Means berlangsung melalui langkah-langkah tersebut untuk membentuk kelompok-kelompok data yang memiliki kesamaan karakteristik dalam cluster-cluster yang berbeda.

Untuk melakukan pemrosesan data pada titik-titik pusat setiap klaster, digunakan teori jarak Euclidean dengan menggunakan rumus berikut.

$$D(i,j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

Rumus 2. 3 Euclidean

Keterangan:

$D(i,j)$: Jarak antara data ke-i dengan pusat klaster j.

X_{ki} : Data pada atribut ke-k dari objek ke-i.

X_{kj} : Pusat dari titik ke atribut j menuju ke atribut k.

2.1.7.3 Algoritma Apriori

Algoritma Apriori adalah algoritma populer dalam teknik penambangan data (data mining) yang bertujuan mencari kombinasi item set dengan frekuensi tertentu sesuai kriteria yang diinginkan. Dengan demikian, hasil algoritma ini membantu manajemen dalam pengambilan keputusan dengan menggunakan teknik association rule untuk mencari frequent itemset berdasarkan pengetahuan sebelumnya tentang itemset yang sering muncul (Djamaludin et al., 2017).

Untuk menemukan kombinasi item yang memenuhi batas minimum nilai support di dalam basis data, kita menggunakan presentase nilai support sebuah item.

Support menggambarkan seberapa sering kombinasi item tertentu muncul dalam database, dan dapat dihitung menggunakan rumus berikut.

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

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

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

Rumus 2. 4 Persamaan nilai support

Setelah berhasil mengidentifikasi semua pola frekuensi tinggi, langkah selanjutnya ialah mencari aturan asosiasi yang memenuhi persyaratan minimum untuk tingkat kepercayaan (*confidence*). Hal ini dilaksanakan dengan menghitung nilai kepercayaan aturan asosiatif A U B menggunakan rumus khusus.

$$\text{Confidence} = P(B|A) = \sum \frac{\text{Transaksi mengandung } A \text{ dan } B}{\text{transaksi}} \quad \text{Rumus 2. 5 Confidence}$$

2.1.7.4 Algoritma K-Nearesst Neighbor

Algoritma ini kerap disebut KNN (K-Nearest Neighbors). KNN adalah metode sederhana namun sangat akurat yang digunakan untuk mengklasifikasikan objek baru berdasarkan jarak terdekat dari instance query ke sampel latihan. Dalam algoritma KNN, data baru yang belum diketahui kelasnya akan diklasifikasikan dengan cara memilih sejumlah data yang terletak paling dekat dari data baru tersebut. Algoritma KNN memanfaatkan metode klasifikasi berdasarkan ketetanggaan untuk melakukan prediksi terhadap data baru (Yolanda & Fahmi, 2021). Prosedur implementasi K-Nearest Neighbors (KNN) melibatkan langkah-langkah kritis yang harus diikuti untuk melakukan klasifikasi data. Pertama-tama,

penetapan parameter K merupakan langkah awal yang penting. Parameter ini mengacu pada jumlah tetangga terdekat yang akan digunakan dalam proses klasifikasi. Selanjutnya, pada tahap kedua, jarak antara data baru yang akan diklasifikasikan dan seluruh data pelatihan dihitung dengan menggunakan rumus Euclidean. Proses perhitungan jarak ini memungkinkan penemuan tetangga terdekat dari data baru, yang kemudian akan menjadi dasar untuk prediksi. Dengan mengikuti langkah-langkah KNN ini, prediksi dapat dihasilkan dengan tingkat akurasi yang signifikan, dan metode ini menjadi salah satu pilihan yang efektif dalam klasifikasi data dalam konteks analisis akademik dan penambangan data.

$$d_i = \sqrt{\sum_{1=i}^k (X_i - X_j)^2} \quad \text{Rumus 2. 6 Euclidean}$$

Keterangan:

X1: Data sampel atau data pelatihan

X2: Data tes atau data pengujian.

D: Perbedaan antara dua titik

K: Atribut/Data memiliki dimensi.

Langkah-langkah yang terjadi dalam proses tersebut dapat dijelaskan sebagai berikut. Urutkan jarak data baru. Setelah itu, cari tetangga terdekat dengan menghitung jarak minimum ke-K, periksa kelas dari tetangga terdekat, dan gunakan mayoritas kelas tetangga terdekat sebagai nilai prediksi untuk data baru.

2.1.7.5 Algoritma Naive bayes

Pengklasifikasian dengan mengimplementasikan metode probabilitas serta statistik ini adalah suatu cara untuk memprediksi peluang kejadian di masa depan berlandaskan pengalaman dari masa lalu. Metode ini dikenal sebagai Teorema Bayes. Dalam penerapannya, teorema Bayes digabungkan dengan pendekatan "Naive," yang mengasumsikan bahwa atribut-atribut yang terlibat dalam klasifikasi adalah saling bebas atau tidak memiliki ketergantungan antara satu sama lain (Putro et al., 2020). Adapun rumus yang dapat digunakan sebagai berikut:

$$P(C_i|X) = P(X|C_i) \cdot \frac{P(C_i)}{P(X)} \quad \text{Rumus 2. 7 Teorema bayes}$$

Keterangan:

X: Data yang memiliki kelas yang belum diketahui.

C_i: Hipotesis tentang data X adalah bahwa itu adalah suatu kelas spesifik.

P(C_i|X): Probabilitas hipotesis C_i yang didasarkan pada kondisi X (disebut juga Probabilitas Posteriori).

P(C_i): Probabilitas dari suatu hipotesis.

P(X|C_i): Probabilitas dari X yang bergantung pada kondisi dari hipotesis C_i.

P(X): Probabilitas terjadinya X.

2.1.7.6 Algoritma *FP-GROWTH*

FP-GROWTH adalah sebuah algoritma yang berasal dari pengembangan algoritma Apriori. Algoritma ini digunakan untuk mengidentifikasi himpunan data yang sering muncul (frequent item set) dalam sebuah kumpulan data. FP-GROWTH menggunakan struktur data berbentuk pohon yang disebut Frequent

Pattern Tree (FP-Tree) untuk memproses data transaksi (Setyo et al., 2019). Tahapan-tahapan dalam metode *FP-Growth* adalah sebagai berikut. Tahap pertama dalam algoritma *FP-growth* adalah pembangkitan *Conditional Pattern Base*, di mana *subdatabase* yang terdiri dari *prefix path* dan *suffix pattern* dibuat melalui penggunaan *FP-tree* yang telah dibangun sebelumnya. Selanjutnya, dalam tahap pembangkitan *Conditional FP-Tree*, dilakukan penjumlahan dukungan untuk setiap item dalam setiap pattern kondisional dasar. Item-item yang memiliki dukungan lebih besar atau sama dengan minimum dukungan akan dimasukkan ke dalam *FP-tree* kondisional. Setelah itu, tahap terakhir adalah pencarian *frequent item set*, di mana berbagai algoritma data mining diterapkan untuk menganalisis pola pembelian konsumen dalam *Conditional FP-tree*. Proses ini melibatkan kombinasi item untuk mendapatkan set item yang sering muncul. Jika *Conditional FP-tree* merupakan lintasan tunggal, maka pembangkitan *FP-growth* akan dilakukan secara rekursif guna mencari *frequent item set* yang lebih kompleks.

Algoritma *FP-Growth* memiliki karakteristik unik yang efektif dalam menemukan *frequent itemset* dari dataset transaksi melalui penggunaan struktur data *FP-Tree*. Dengan *FP-Tree*, algoritma ini dapat langsung mengekstrak *frequent itemset* dari data transaksi tanpa memerlukan langkah kandidat *itemset generation* seperti pada algoritma Apriori. Prosesnya melibatkan langkah-langkah seperti menentukan Minimum Support, membentuk *Header Frequent Itemset*, dan melakukan pembentukan *FP-Tree* dengan pengompresian data. Selanjutnya, *Conditional Pattern* dibuat dari *FP-Tree* dan digunakan untuk menentukan Frequent Item-set dengan bantuan algoritma rekursif atau *FP-Growth* itu sendiri.

Dengan pendekatan efisien dan minimnya langkah-langkah yang memerlukan pembentukan kandidat itemset, algoritma *FP-Growth* mampu mengatasi masalah itemset yang besar dan kompleks secara cepat dan efektif, memberikan solusi yang handal untuk analisis data transaksi.

2.1.8 Association Rule Mining

Association rule mining, yang juga dikenal sebagai analisis asosiasi, adalah salah satu teknik dalam data mining yang digunakan untuk menemukan aturan asosiasi. dari analisis dari suatu pembelian sparepart yang telah diketahui beberapa besar kemungkinan seorang pelanggan membeli oli motor dan busi motor secara bersamaan. Analisis asosiasi juga dikenal dengan market based analisis yang diketahui sebagai teknik dasar pada data mining. Salah satu fase dari analisis asosiasi yang dikenal sebagai *frequent pattern mining* adalah tahap analisis pola frekuensi tinggi.

Association Rule Mining adalah teknik data mining yang mencari pola-pola asosiasi relevan dalam data transaksi atau itemset. Algoritma Apriori sering digunakan untuk mengidentifikasi item-item yang sering muncul bersama dalam aturan asosiasi. Dengan teknik ini, organisasi dapat menemukan pola-pola menarik dan bermanfaat dari data transaksi mereka, memberikan kontribusi penting dalam pengambilan keputusan yang lebih efisien. Cara untuk memperoleh nilai dukungan dari suatu item A adalah melalui rumus berikut ini:

$$\text{Support } A = \frac{\text{Jumlah transaksi yang mengandung item } A}{\text{total transaksi}};$$

Rumus 2. 8 Nilai support dari suatu item A

Setelah itu, untuk memperoleh nilai dukungan dari dua item dapat dilakukan menggunakan rumus berikut:

$$\begin{aligned} \text{Support}(A, B) &= P(A \cap B) \\ &= \frac{\text{jumlah transaksi yang mengandung } A \text{ dan } B}{\text{total transaksi}}; \end{aligned}$$

Rumus 2. 9 Nilai support dari dua item

Setelah memperoleh semua item yang sering muncul (*frequent item*) dan himpunan item besar (*large item set*), kita bisa mencari nilai kepercayaan minimum (*mincof*) dengan menggunakan rumus berikut:

$$\begin{aligned} \text{Confidence}(A \rightarrow B) &= P(A|B) \\ &= \frac{\text{Jumlah transaksi yang mengandung } A \text{ dan } B}{\text{jumlah transaksi yang mengandung } A}; \end{aligned}$$

Rumus 2. 10 Minimum confidence (mincof)

2.2 *Software* Pendukung

2.2.1 Pengertian *Rapid Miner*

RapidMiner adalah perangkat lunak analisis data yang sangat efisien dan populer. Penggunaannya melibatkan prinsip-prinsip analisis data, seperti eksplorasi, pemrosesan, dan transformasi data untuk memahami struktur dan informasi yang terkandung dalam set data. Selain itu, *RapidMiner* juga memberikan dukungan penuh untuk pembelajaran mesin dengan berbagai teknik seperti klasifikasi, regresi, dan clustering. Paradigma eksperimen yang digunakan membantu pengguna untuk menciptakan proses analisis yang mudah dipahami dan diulang. Kelebihan lainnya adalah lingkungan visual yang memungkinkan pengguna membangun alur kerja analisis dengan menggabungkan blok-blok

fungsional, sehingga membantu dalam membuat analisis data yang intuitif. Integrasi data dan proses yang ditawarkan RapidMiner juga sangat berguna dalam pemrosesan data yang lebih komprehensif dan mendalam (Pranata & Utomo, 2020).



Gambar 2. 2 Rapid Minner
Sumber: (Pranata & Utomo, 2020).

2.2.2 Fitur-fitur *Rapid Minner*

RapidMiner menyediakan beberapa fitur dalam menangani pekerjaan yang terkait dengan *data mining*. Pertama, perangkat ini menawarkan visualisasi yang canggih, seperti diagram histogram, diagram pohon, dan diagram sebar 3D, yang memudahkan pengguna dalam memahami dan menganalisis data secara efisien. Selanjutnya, *RapidMiner* memiliki banyak pilihan *plugin* termasuk *plugin* teks yang memungkinkan analisis teks mining, memberikan fleksibilitas dalam mengolah data teks secara lebih mendalam. Selain itu, *RapidMiner* menyediakan teknologi data mining dan machine learning yang lengkap, termasuk ETL (*Extract, Transform, Load*) pra-pemrosesan data, visualisasi data, pemodelan, dan evaluasi. Proses penambangan data di tulis dalam format XML dan terdiri dari operator nestable yang dapat disusun secara intuitif melalui GUI. Terakhir, *RapidMiner* juga mengintegrasikan proyek data mining Weka dan statistika R, sehingga pengguna dapat memanfaatkan kekuatan dari kedua platform ini dalam analisis data mereka.

Dengan fitur-fitur yang luar biasa ini, RapidMiner menjadi pilihan utama bagi banyak profesional dan organisasi untuk menggali wawasan berharga dari data mereka.

2.2.3 Microsoft Excel

Microsoft Excel adalah perangkat lunak pengolahan data dengan fitur lembar kerja, rumus, grafik, dan analisis data, serta dukungan bahasa pemrograman VBA untuk otomatisasi tugas. Berguna dalam berbagai bidang seperti akuntansi, keuangan, sains, dan bisnis (Pratama and Arumi 2022).



Gambar 2.3 *Microsoft Excel*
Sumber: (Indra Borman and Wati 2020).

2.3 Penelitian Terdahulu

Studi sebelumnya atau penelitian terdahulu adalah jurnal penelitian yang telah dilaksanakan oleh para peneliti sebelumnya dan kini digunakan dalam penelitian ini sebagai referensi bagi peneliti.

1. Berdasarkan penelitian yang telah dilakukan terlebih dahulu oleh (Pranata & Utomo, 2020) dengan judul “Penerapan Data Mining Algoritma FP-Growth Untuk Persediaan Sparepart Pada Bengkel Motor” Bulletin Of Information Teknologi (BIT) ISSN 2722-0524 Hasil penelitian menyimpulkan bahwa

dari pengujian yang dilakukan, didapatkan nilai tertinggi 8.9 untuk spartpart x dan y, sedangkan nilai terendah untuk s dan t, berdasarkan proses *fp-growth* pada data bengkel *service*.

2. Berdasarkan penelitian yang telah dilakukan terlebih dahulu oleh (Lintas, 2017) dengan judul “Implementasi Algoritma FP-Growth Dalam Penentuan Pola Hubungan Kecelakaan Lalu Lintas” Jurnal Sistem Informasi 2/13 (2017) DOI:<http://dx.doi.org/10.21609/jsi.v13i2.551> Penelitian menyimpulkan bahwa berhasil mencapai tujuan yang diharapkan, yaitu membantu pihak terkait dalam membuat keputusan untuk mengurangi kecelakaan.
3. Berdasarkan penelitian yang telah dilakukan terlebih dahulu oleh (Fp-growth et al., 2022) dengan judul “Implementasi Data Mining Dalam Menganalisa Pola Penjualan Roti Menggunakan Algoritma FP-Growth” Jurnal Sistem Informasi TGD Vol 1no 3 mei 2022 ISSN 2828-1004 Hasil penelitian menyimpulkan bahwa untuk menerapkan data mining dengan metode FP-Growth, dilakukan dengan membuat aplikasi desktop untuk melakukan proses perhitungan, dan juga merancang basis data yang relevan.
4. Berdasarkan penelitian yang telah dilakukan terlebih dahulu oleh (Jhoni & Medan, 2015) dengan judul “Penerapan Data Mining Menggunakan Algoritma FP-Tree dan FP-Growth Pada Data Transaksi Penjualan Obat ” Seminar Nasional Teknologi Informasi dan Komunikasi (SNASTIKOM 2015) Penelitian menyimpulkan bahwa implementasi dapat dilakukan dengan memanfaatkan database penjualan obat karena mampu menemukan pola

kombinasi item set yang akan membantu dalam mengembangkan strategi penjualan kepada konsumen.

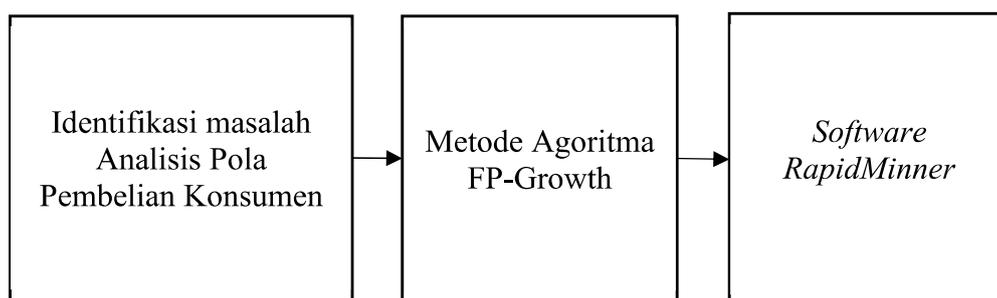
5. Berdasarkan penelitian yang telah dilakukan terlebih dahulu oleh (Setyo et al., 2019) dengan judul “Implementasi Data Mining Pada Penjualan Produk Di CV Cahaya Setya Menggunakan Algoritma FP-Growth” Jurnal Petir vol 12 no 1 maret 2019 ISSN 1978-9262 Hasil penelitian menyimpulkan bahwa ketika nilai minimum support semakin kecil, maka akan dihasilkan lebih banyak aturan asosiasi, dan sebaliknya, jika nilai minimum support semakin besar, maka aturan asosiasi yang dihasilkan akan lebih sedikit.
6. Berdasarkan penelitian yang telah dilakukan terlebih dahulu oleh (Kasus et al., 2012) dengan judul “Penerapan Metode Data Mining Market Basket Analisis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori dan Frequent Pattern Growth (FP_Growth) ” Jurnal Telematika Mkom vol 4 no 1 ISSN 2085-725X Hasil penelitian menyimpulkan bahwa berbagai algoritma menghasilkan perbedaan dalam jumlah aturan asosiasi yang dihasilkan.
7. Berdasarkan penelitian yang telah dilakukan terlebih dahulu oleh (Maulidiya et al., 2020) dengan judul “Asosiasi Data Mining Menggunakan Algoritma Apriori dan FP-Growth Sebagai Dasar Pertimbangan Penentuan Paket Sembako” Prociding SENDIU 2020 Hasil penelitian menyimpulkan bahwa algoritma FP-Growth memiliki tingkat akurasi yang jauh lebih tinggi daripada algoritma Apriori, yakni mencapai 284%, yang artinya tiga kali lebih besar daripada tingkat akurasi algoritma Apriori. Selain itu, algoritma FP-

Growth juga mampu membentuk kombinasi item tertinggi hingga 3 kombinasi item.

8. Berdasarkan penelitian yang telah dilakukan oleh (Wijaya & Fauzi, 2020) dengan judul "Data Mining Algoritma C4.5 Untuk Memprediksi Penjualan Baterai Di PT Varta Microbattery Indonesia" Jurnal Comasie vol 3 no 2 (2020) Kesimpulan dari penelitian menunjukkan bahwa menggunakan metode algoritma C4.5 dengan pohon keputusan dapat memberikan informasi tentang prediksi penjualan baterai melalui representasi aturan yang dihasilkan.

2.4 Kerangka Pemikiran

Kerangka pemikiran adalah konsepsi dari inti permasalahan yang akan dipelajari, serta menggambarkan langkah-langkah dalam proses awal hingga akhir penelitian dalam bentuk desain alur urutan penyelesaiannya. Di bawah ini terdapat kerangka pemikiran yang telah disusun oleh peneliti.



Gambar 2. 4 Kerangka Pemikiran
Sumber: (Peneliti, 2023)

Data yang di input berupad masalah pada analisis pola pembelian konsumen. Data penelitian yang telah dikumpul akan dilakukan perhitungan penyelesaian

menggunakan algoritma FP-Growth. Hasil yang diharapkan akan dilakukan pengujian menggunakan rapidminer sehingga menghasilkan spartpare mana saja yang memiliki nilai jual tinggi berdasarkan pola pembelian konsumen