

BAB II

LANDASAN TEORI

2.1 Definisi Penjualan

Menurut (Darmaastawan 2023) Penjualan adalah bagaimana menciptakan hubungan jangka panjang dengan pelanggan melalui produk atau jasa. Penjualan adalah proses antara penjual dan pembeli untuk menukar barang atau jasa dengan uang atau nilai lainnya. Proses ini mencakup berbagai tindakan, seperti mempromosikan barang atau jasa, mendiskusikan harga, menyelesaikan transaksi, dan mengirimkan barang atau jasa kepada pelanggan.

Dalam konteks bisnis, penjualan itu sangat penting bagi perusahaan. Proses penjualan yang efektif memerlukan pemahaman yang mendalam tentang kebutuhan dan keinginan pelanggan serta kemampuan untuk memberi tahu mereka tentang nilai barang atau jasa.

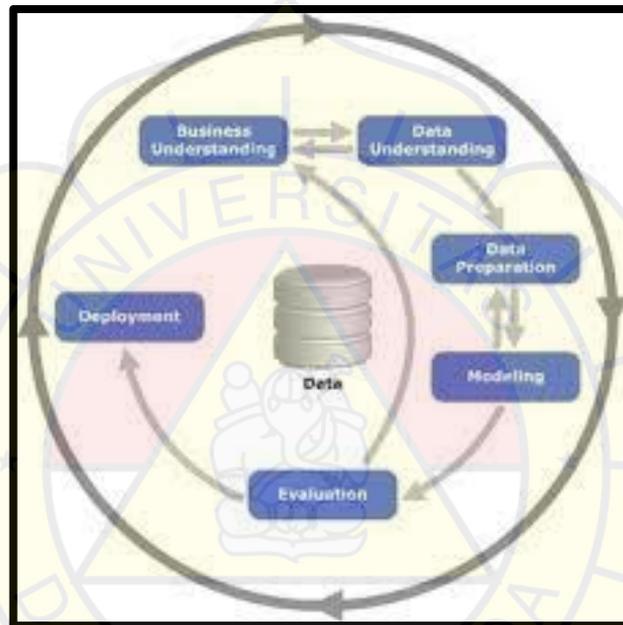
2.2 Data Mining

Menurut (Napila 2023) Data mining adalah proses penggalian dan pencarian pengetahuan dan informasi yang bermanfaat dengan menggunakan algoritma atau metode teknik tertentu sesuai dengan pengetahuan atau informasi yang dicari.

Data mining juga merupakan proses yang digunakan untuk mengekstraksi pola, informasi, atau pengetahuan yang bermanfaat atau tidak terstruktur dari kumpulan data yang sangat besar. Tujuan utama dari proses ini adalah untuk menemukan tren, pola, atau hubungan yang bermanfaat dalam data yang dapat membantu membuat keputusan yang lebih baik dalam berbagai bidang.

2.2.1 CRISP-DM

CRISP-DM merupakan suatu tata cara yang dijadikan sebagai proses standarisasi dalam proses data mining yang didirikan oleh kumpulan perusahaan komite eropa pada tahun 1996. CRISP-DM menyajikan standar proses dalam pengolahan data mining sebagai siasat untuk dapat memecahkan suatu masalah dari perspektif bisnis atau bagi penelitian secara umum.(Butar and Kurnia 2023)



Gambar 2.1. CRISP-DM

Dalam proses pengolahan penelitian menggunakan data mining, CRISP-DM terdiri dari enam tahapan dan proses. Tahapan tersebut terdiri dari yaitu:

1. *Business Understanding* (Pemahaman Bisnis)

Pemahaman bisnis atau *Business Understanding* merupakan tahap yang sangat penting untuk memahami jenis pekerjaan data mining yang akan dilakukan dan kebutuhan dari sudut pandang bisnis. Salah satu tugasnya adalah memahami situasi bisnis, menentukan tujuan bisnis, menentukan tujuan data mining, dan membuat strategi dan jadwal penelitian.

2. *Data Understanding* (Pemahaman Data)

Proses ini dimulai dengan proses pengumpulan data yang akan digunakan. Selanjutnya, tahap pemahaman data akan dilakukan untuk mendapatkan pemahaman yang lebih mendalam tentang data yang akan digunakan.

3. *Data Preparation* (Pengolahan Data)

Pada tahapan proses pengolahan data berkaitan dengan menghapus atau menghilangkan beberapa atribut yang tidak digunakan dan tidak dipakai.

4. *Modeling* (Pemodelan)

Ini adalah bagian dari desain penelitian, mulai dari persiapan data, memilih metode pemrosesan data, dan menggabungkan standar dan kriteria metode pemrosesan data dengan nilai terbaik dan ideal.

5. *Evaluation* (Evaluasi)

Pada langkah ini, hasil dari tahap modeling sebelumnya dievaluasi. Ini dilakukan dengan tujuan untuk menyesuaikan hasil akhir dari tahap modeling dengan tujuan awal.

6. *Deployment* (Penyebaran)

Bagian dari proses membuat laporan tentang hasil pengolahan data mining. Laporan akhir adalah pengetahuan yang diperoleh atau didapat dari proses pengolahan data mining dan dipresentasikan dalam bentuk grafik atau deskripsi sehingga orang lain dapat memahami dan memahaminya.

2.2.2 Pengelompokan Data Mining

Adapun metode-metode pada data mining yang umum digunakan, diantaranya:

1. Klasifikasi

Teknik ini digunakan untuk mengklasifikasikan data ke dalam kategori atau kelas yang sudah ditentukan berdasarkan pola atau karakteristik tertentu.

2. Klustering

Klustering digunakan untuk menganalisis data kedalam kelompok atau klaster berdasarkan kesamaan karakteristik.

3. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategorik. Model dibangun menggunakan baris data (record) lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi.

4. Prediksi

Teknik atau metode yang bertujuan untuk membangun model yang dapat digunakan untuk memprediksi nilai atau perilaku di masa depan berdasarkan data historis yang ada.

5. Asosiasi

Merupakan metode atau teknik yang digunakan untuk mengklasifikasikan data ke dalam kategori atau kelas yang sudah ditentukan berdasarkan pola atau karakteristik tertentu.

2.2.3 *Association Rule*

Menurut (Patandean, Pasinggi, and Paelleng 2023) Aturan asosiasi atau *association rule* adalah salah satu pendekatan data mining yang dapat digunakan untuk menemukan aturan asosiatif item. Aturan asosiasi ini dibuat dengan mengidentifikasi hubungan yang paling signifikan dari pola data yang sering terjadi (*frequent pattern*) dengan menggunakan parameter support dan confidence.

Support adalah indikasi metrik yang mengungkapkan seberapa sering item database muncul. Sementara confidence menunjukkan seberapa sering suatu pernyataan tersebut akurat.

Pada transaksi yang terdapat item X terdapat kemungkinan ada item Y juga didalamnya, dinotasikan $X \rightarrow Y$, dimana X dan Y adalah disjoint itemset, dinotasikan $X \cap Y$. kumpulan dari transaksi-transaksi ini disebut dengan itemset, yang dinotasikan dengan I_k ($k= 1, 2, \dots m$). jika terdapat itemset yang mempunyai item sebanyak k, maka disebut k-itemset.

Association rule ini nantinya menghasilkan rules yang menentukan seberapa besar hubungan antar X dan Y tadi, dan diperlukan ukuran untuk rules ini yakni support dan confidence. Support merupakan kemungkinan X dan Y muncul bersamaan yang dinotasaikan:

$$\text{Support } (X \rightarrow Y) = \frac{\text{Jumlah transaksi yang mengandung X dan Y}}{\text{Jumlah transaksi}}$$

Sedangkan confidence merupakan kemungkinan munculnya Y ketika X juga muncul, dinotasikan:

$$\text{Confidence } (X \rightarrow Y) = \frac{\text{Jumlah transaksi yang mengandung X dan Y}}{\text{Jumlah transaksi yang mengandung X}}$$

2.3 Algoritma Apriori

Apriori adalah salah satu algoritma yang melakukan pencarian *frequent itemset* dengan menggunakan teknik *association rule*. Algoritma apriori adalah suatu metode untuk mencari pola hubungan antar satu atau lebih item dalam suatu dataset.

Dalam data mining, algoritma Apriori digunakan untuk menemukan hubungan atau asosiasi antara item dalam kumpulan data transaksi. Ini adalah salah satu metode yang paling umum untuk menemukan aturan hubungan dalam data. Apriori memakai pendekatan iteratif menggunakan *level-wise search* dimana k-itemset digunakan dalam mencari (k+1)-itemset. iterasi i menghitung seluruh deretan data i (deretan yang mengandung elemen i) yang seringkali timbul. Setiap iterasi terdiri berasal dari dua langkah yaitu *candidate generation* (penentuan kandidat) serta *candidate counting and selection* (pemilihan serta penghitungan kandidat).

2.3.1 Tahapan Pengolahan Apriori

Adapun proses utama yang dilakukan dalam algoritma Apriori, antara lain sebagai berikut (Soepriyono and Triayudi 2023):

1. *Join Step* (Penggabungan) Proses ini akan mengombinasikan setiap item dengan item lainnya sampai tidak terbentuk kombinasi lagi.
2. *Prune Step* (Pemangkasan) Hasil dari kombinasi item akan dipangkas dengan minimum support yang telah ditentukan oleh user.

Menurut (Soepriyono and Triayudi 2023) dalam langkah awal pada Algoritma Apriori, kita mencari frequent itemset dengan menganalisis pola frekuensi tertinggi. Hal ini dilakukan dengan mencari kombinasi item yang memenuhi nilai support minimum dalam basis data atau dataset.

Support dari tiap kandidat k-itemset didapat menghitung jumlah transaksi yang memuat semua item di dalam kandidat k-itemset tersebut. Ini adalah juga ciri dari algoritma apriori yaitu diperlukan penghitungan dengan scan seluruh database

sebanyak k-itemset terpanjang. Untuk menghitung nilai support dengan satu item, rumus yang digunakan adalah:

$$\text{Support } A = \frac{\text{Jumlah Transaksi mengandung } A}{\text{Jumlah Transaksi}} \times 100\%$$

Lalu untuk rumus menghitung nilai support untuk 2 itemset menggunakan rumus dibawah ini:

$$\text{Support } (A, B) = \frac{\text{Jumlah Transaksi mengandung } A,B}{\text{Jumlah Transaksi}} \times 100\%$$

Setelah semua pola frekuensi tinggi telah ditemukan, langkah berikutnya adalah membuat aturan asosiasi untuk memastikan bahwa tingkat confidence terpenuhi. Ini dilakukan dengan menghitung nilai confidence untuk setiap aturan asosiasi, yang berarti "jika A maka B". (Juliano, Rasim, and Sugiyatno 2022)

2.4 Algoritma Eclat

Menurut (Lisnawita and Devega 2018) Dalam data mining, algoritma Eclat (*Equivalence Class Transformation*) digunakan untuk menemukan aturan asosiasi antara item dalam dataset transaksi. Algoritma ini mirip dengan algoritma Apriori, tetapi menggunakan metode yang berbeda untuk menemukan itemset yang signifikan.

2.4.1 Tahapan Pengolahan Algoritma Eclat

Menurut (Sudarsono 2019) pada dasarnya algoritma ECLAT melakukan pencarian secara *depth-first search* pada database dengan tata letak vertikal, apabila database berbentuk horizontal maka terlebih dahulu harus dikonversikan ke dalam bentuk vertical.

Algoritma Eclat tidak terlalu kompleks secara umum, tetapi implementasinya dapat lebih rumit dalam praktiknya tergantung pada bahasa pemrograman atau alat yang digunakan. Secara umum, alur algoritma Eclat dapat digambarkan sebagai berikut:

1. Menemukan itemset kandidat

Pada tahap ini, algoritma mencari semua set item tunggal yang muncul dalam transaksi dan menggabungkannya untuk membentuk itemset yang lebih besar.

2. Menghitung Support

Setelah itemset kandidat ditemukan, langkah selanjutnya adalah menghitung dukungan untuk masing-masing itemset. Dukungan adalah frekuensi relatif itemset dalam data transaksi.

3. Menggali itemset yang Sering Muncul

Itemset yang memiliki dukungan di atas batas minimum yang ditentukan, atau batas dukungan, dianggap sebagai itemset yang sering muncul dalam kumpulan data transaksional.

Rumus matematis untuk menghitung support suatu itemset dapat direpresentasikan sebagai berikut:

$$\text{Support } X = \frac{\text{Jumlah Transaksi mengandung } X}{\text{Jumlah Transaksi}}$$

Di mana:

1. X adalah itemset yang sedang dievaluasi.
2. Jumlah transaksi yang berisi X adalah jumlah transaksi yang mengandung semua item dalam X.
3. Total transaksi adalah total jumlah transaksi dalam kumpulan data.

2.5 RStudio

RStudio adalah lingkungan pengembangan terintegrasi yang populer untuk bahasa pemrograman R yang memungkinkan proses pengembangan, analisis, dan visualisasi data. Dengan menggunakan RStudio, pengguna dapat melakukan serangkaian tugas analisis data, termasuk klasifikasi, klusterisasi, regresi, asosiasi, dan lain-lain, tanpa perlu menulis kode pemrograman.

RStudio juga menawarkan antarmuka pengguna yang ramah dan intuitif, yang memudahkan pengguna dalam menulis, menjalankan, dan mengelola skrip R. Antarmuka ini terbagi menjadi beberapa panel yang menampilkan editor kode, konsol R, tampilan data, dan jendela plot grafik secara terorganisir.

2.6 Pengembangan Sistem

Pengembangan sistem adalah salah satu tahap krusial dalam sebuah penelitian yang menerapkan metode analisis data mining seperti algoritma Apriori dan Eclat. Pada dasarnya, tahap pengembangan sistem meliputi beberapa aspek penting yang perlu diperhatikan dengan seksama. Pertama-tama, kita perlu mempersiapkan infrastruktur perangkat lunak dan perangkat keras yang dibutuhkan untuk menjalankan algoritma-algoritma yang akan digunakan. Ini termasuk pemilihan bahasa pemrograman, *framework* atau *library* yang sesuai, serta pemilihan platform komputasi yang mampu menangani ukuran dan kompleksitas data yang dimiliki.

Selanjutnya, langkah penting adalah mempersiapkan data yang akan digunakan dalam analisis. Data penjualan mungkin tersimpan dalam format yang beragam, seperti database SQL, file CSV, atau bahkan dokumen teks. Oleh karena

itu, dalam tahap ini, perlu mengembangkan skrip atau rutinitas untuk mengekstraksi dan membersihkan data tersebut agar siap untuk diolah oleh algoritma Apriori dan Eclat.

2.6.1 PHP

Untuk memanfaatkan algoritma Apriori dan Eclat, diperlukan bahasa pemrograman yang memiliki kemampuan untuk mengelola manipulasi data secara efektif dan efisien. PHP atau singkatan dari *Hypertext Preprocessor* adalah bahasa pemrograman open source yang umum digunakan untuk pengembangan web dan awalnya dibuat untuk membuat halaman web dinamis. Namun, seiring waktu, PHP telah berkembang menjadi bahasa pemrograman yang kuat untuk berbagai jenis pengembangan perangkat lunak.

Dengan demikian, penggunaan PHP sebagai bahasa pemrograman untuk mengimplementasikan algoritma data mining dapat memberikan keuntungan dalam hal kemudahan penggunaan, ketersediaan sumber daya, pustaka yang kuat, dan kinerja yang cukup baik. Oleh karena itu, PHP adalah pilihan yang tepat untuk mengembangkan sistem analisis menggunakan algoritma data mining.

2.6.2 XAMPP

XAMPP (X(Windows/Linux) Apache PHP dan Perl) merupakan paket server web PHP dan database MySQL yang paling populer di kalangan pengembangan web dengan menggunakan PHP dan MySQL sebagai databasenya. Paket XAMPP, sesuai dengan kepanjangannya, X yang berarti Windows atau Linux, pengguna bisa memilih paket yang diinginkan untuk Windows atau Linux, pengguna bisa memilih paket yang diinginkan untuk Windows atau Linux. XAMPP termasuk paket server

yang paling mudah untuk digunakan sebagai paket untuk pengembangan aplikasi web.

2.6.3 Usecase Diagram

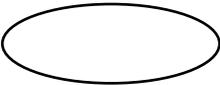
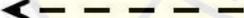
Usecase diagram adalah sebuah skenario yang menjelaskan bagaimana suatu sistem atau aplikasi akan digunakan oleh pengguna untuk mencapai tujuan tertentu. Usecase biasanya menjelaskan alur interaksi antara pengguna dan sistem, termasuk masukan yang diberikan oleh pengguna, bagaimana sistem menanggapi masukan tersebut, dan hasil atau output yang dihasilkan oleh sistem.

Usecase juga dapat mencakup berbagai alternatif atau cabang dari alur utama yang menjelaskan bagaimana sistem menangani berbagai situasi atau pemrosesan yang kompleks. Tujuan utama dari penggunaan *usecase* adalah untuk mempelajari lebih lanjut tentang kebutuhan dan keinginan pengguna.

Berikut ini adalah simbol-simbol diagram usecase, seperti yang terlihat pada tabel dibawah ini:

Tabel 2.1 Simbol-simbol Usecase Diagram

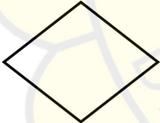
SIMBOL	NAMA	KETERANGAN
	Aktor	Seseorang atau sesuatu yang berinteraksi dengan sistem yang sedang kita kembangkan.

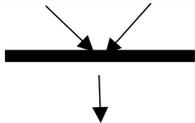
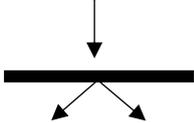
	<p>Usecase</p>	<p>Fungsionalitas yang disediakan sistem sebagai unit-unit yang saling bertukar pesan antar unit dan aktor</p>
	<p>Asosiasi</p>	<p>Komunikasi antar aktor dan use case yang berpartisipasi</p>
<p><<Include>></p> 	<p>Include Relationship</p>	<p>Relasi cakupan memungkinkan suatu use case untuk menggunakan fungsionalitas yang disediakan oleh use case yang lainnya.</p>
<p><<Extend>></p> 	<p>Extend Relationship</p>	<p>Memungkinkan suatu use case memiliki kemungkinan untuk memperluas fungsional yang disediakan use case yang lainnya.</p>

2.6.4 Activity Diagram

Activity diagram merupakan gambarkan proses bisnis dan urutan aktivitas dalam sebuah proses. Diagram ini sangat mirip dengan *flowchart* karena memodelkan *workflow* dari satu aktivitas ke aktivitas lainnya atau dari aktivitas ke status. *Activity diagram* juga bermanfaat untuk menggambarkan *parallel behaviour* atau menggambarkan interaksi antara beberapa usecase.

Tabel 2.2 Simbol-simbol Activity Diagram

SIMBOL	NAMA	KETERANGAN
	Status Awal	Status awal aktivitas system
	Aktivitas	Menggambarkan sebuah kegiatan atau tugas yang perlu dilakukan
	Percabangan / <i>Decision</i>	Asosiasi percabangan dimana jika ada pilihan aktivitas lebih dari satu
	Penggabungan / <i>Join</i>	Asosiasi penggabungan dimana lebih dari satu aktivitas digabungkan menjadi satu
	Status Akhir	Status akhir aktivitas system

	<p style="text-align: center;"><i>Join</i></p>	<p>Digunakan untuk menunjukkan kegiatan yang digabungkan</p>
	<p style="text-align: center;"><i>Fork</i></p>	<p>Digunakan untuk menunjukkan kegiatan yang dilakukan secara parallel</p>

2.7 Kajian Penelitian Terdahulu

Berdasarkan teori dan penelitian terdahulu, maka perlu menggunakan data dan referensi topik penelitian untuk memudahkan penyusunan laporan penelitian. Data pendukung yang perlu diperhatikan dalam penyusunan laporan penelitian ini adalah data-data pendukung yang relevan dengan penelitian sebelumnya dan memuat apa yang dibahas dalam penelitian ini. Sebagai bahan pertimbangan dalam penelitian ini telah disertakan beberapa hasil penelitian sebelumnya oleh beberapa peneliti ternama, antara lain:

Pertama, penelitian (Darmaastawan 2023) yang berjudul Analisis Keranjang Belanja Menggunakan Algoritma Apriori untuk Menentukan Tata Letak Barang. Penelitian ini menghasilkan kesimpulan bahwa data mining menggunakan Algoritma Apriori dapat digunakan untuk menemukan informasi tersembunyi pada data transaksi Apotek Mooladhara selama periode Desember 2022 hingga Februari 2023. Data transaksi tersebut berjumlah 581 transaksi dan terdiri dari 270 barang. Proses data mining menggunakan Algoritma Apriori dengan nilai support minimum adalah 2% dan nilai confidence minimum adalah 30% terhadap data transaksi

berhasil membentuk beberapa association rules atau aturan asosiasi yang menyimpan informasi barang yang sering dibeli oleh pelanggan apotek secara bersamaan.

Kedua, penelitian (Napila 2023) berjudul Analisis Metode Market Basket Analysis Terhadap Kegiatan Penjualan Catering Dengan Algoritma Apriori Dan Algoritma Frequent Pattern Growth. Penelitian ini dilakukan dengan membandingkan kedua algoritma, menghasilkan algoritma yang lebih tepat dalam menentukan Association Rule yang digunakan sebagai promosi paket menu yang baru pada Zara Catering. Untuk menghasilkan tingkat akurasi adalah dengan cara menghitung minimum support dan confidence yang dihasilkan.

Ketiga, penelitian (Patandean, Pasinggi, and Palelleng 2023) yang berjudul Market Basket Analysis Menggunakan Algoritma ECLAT (Studi Kasus: Transaksi Penjualan di Minimarket). Penelitian ini menghasilkan kesimpulan bahwa Penggunaan algoritma Eclat dalam menganalisa data transaksi penjualan berhasil diterapkan dalam Market Basket Analysis dan mendapatkan hasil sehingga menjadi sangat efisien dan mempercepat dalam proses perhitungan. Hasil penggunaan algoritma ini dapat diterapkan sebagai salah satu teknik peletakan barang pada Minimarket Trabas agar strategis sehingga menyebabkan berkurangnya barang yang ada di Minimarket Trabas yang kurang laku dan kadaluwarsa sehingga tidak menyebabkan kerugian.