

**BAB II**  
**LANDASAN TEORI**

**1.1 Tinjauan Pustaka**

Pada bagian ini terdapat beberapa penelitian terdahulu sebagai bahan pendukung penelitian yang akan peneliti lakukan, dapat dilihat pada **Tabel 2.1**

Tabel 2. 1 Tinjauan Pustaka

<b>No</b>	<b>Penulis</b>	<b>Judul</b>	<b>Keterangan</b>	<b>Kekurangan dan Kelebihan</b>
1	Sopiyan Ependi, Muhamad Akbar, 2021	Implementasi Data Mining Pada Penjualan Produk Dengan Menggunakan Algoritma Apriori	Penelitian ini bertujuan untuk membantu perusahaan dalam mengambil keputusan yang lebih efektif dalam memprioritaskan pembelian suku cadang. Metode yang digunakan adalah algoritma apriori. Hasilnya membantu perusahaan dalam menganalisis pola transaksi dan memprioritaskan pembelian suku cadang	Kekurangan pada jurnal ini hanya menggunakan 10 data penjualan.
2	Mukhlis Ramadhan, Juniar Hutagalung,	Prediksi Penjualan Spare Part Mobil Daihatsu	Penelitian ini bertujuan untuk menganalisis pola penjualan suku cadang dan	kekurangan pada jurnal tersebut adalah ukuran dataset yang

	Muhammad Dahria, Iskandar Zulkarnain, 2023	Menggunakan Algoritma Apriori	meningkatkan strategi pemasaran bengkel Daihatsu menggunakan algoritma apriori. Metode yang digunakan adalah metode analisis asosiasi dengan menggunakan algoritma apriori. Hasil yang didapat menunjukkan bahwa implementasi sistem prediksi pola penjualan suku cadang menggunakan algoritma apriori dapat mengidentifikasi pola-pola yang kuat antara beberapa kombinasi produk suku cadang mobil Daihatsu.	relatif kecil hanya menggunakan 180 data transaksi. Kelebihan pada jurnal tersebut adalah algoritma apriori terbukti efektif dalam mengidentifikasi pola penjualan yang berasosiasi.
3	Silvi Sintia, Poningsih, Ilham Syahputra Saragih, Anjar Wanto, Irfan Sudahri Damanik, 2019	Penerapan Algoritma Apriori Dalam Memprediksi Hasil Penjualan SparePart PC (Studi Kasus: Toko Sentra Computer)	Penelitian ini bertujuan untuk membantu pemilik toko dalam memprediksi hasil kombinasi penjualan <i>spare part</i> PC dan meningkatkan ketersediaan stok barang menggunakan	Kekurangan pada jurnal ini adalah jumlah data yang digunakan terbatas hanya menggunakan data penjualan dalam kurun waktu satu bulan

			<p>metode algoritma apriori. Hasil penelitian menunjukkan kombinasi pembelian <i>Ram, Lcd, dan Motherboard</i> memiliki tingkat <i>confidence</i> tertinggi sebesar 75%. Sedangkan kombinasi pembelian <i>Hardisk, Motherboard, dan Ram</i> memiliki tingkat <i>confidence</i> tertinggi kedua sebesar 100%.</p>	<p>yaitu bulan desember 2018. Kelebihan pada jurnal ini adalah algoritma apriori terbukti efektif dalam analisis data penjualan dan penggalian aturan asosiasi.</p>
4	(P. Aiswarya dkk., 2020)	Detection and Prediction of Frequent Diseases in India through Association Technique using Apriori Algorithm and Random Forest Regression	<p>Tujuan dari jurnal tersebut adalah untuk mengidentifikasi dan memprediksi penyakit yang sering terjadi di india menggunakan metode algoritma apriori dan <i>random forest regression</i>. Metode yang digunakan adalah algoritma apriori untuk analisis asosiasi dan <i>random forest regression</i> untuk melakukan prediksi. Hasil dari penelitian</p>	<p>Kekurangan dari jurnal ini ialah hasil penelitian tidak dijelaskan secara rinci. Kelebihan dari jurnal ini menggunakan dua algoritma yaitu algoritma apriori untuk melakukan analisis asosiasi dan <i>random forest regression</i> untuk melakukan prediksi.</p>

			tersebut ialah algoritma apriori dan <i>random forest regression</i> dapat digunakan untuk mengidentifikasi dan memprediksi penyakit yang sering terjadi di india.	
5	Adi Sucipto, Hadapiningradja Kusumodestoni, 2022	Implementasi Data Mining Algoritma Apriori Terhadap Penjualan Spare Part Motor di Yamaha Arista Subang	Tujuan dari penelitian ini adalah untuk menganalisis data penjualan dan mengidentifikasi pola dan asosiasi antara berbagai jenis <i>spare part</i> . Metode yang digunakan dalam jurnal ini adalah algoritma apriori untuk menghasilkan <i>itemset</i> dan aturan asosiasi berdasarkan ambang batas minimum <i>support</i> dan <i>confidence</i> . Hasil dari jurnal ini ialah sistem dapat menghasilkan aturan asosiasi menggunakan 206 data penjualan <i>spare part</i> motor dengan batasan	Kekurangan pada jurnal ini adalah jumlah data yang digunakan hanya 206 data transaksi.

			<p>minimum <i>support</i> 15% dan minimum <i>confidence</i> 30%, sehingga menghasilkan 6 pola aturan asosiasi untuk penjualan <i>spare part</i> motor pada bulan juni 2020 di Yamaha Arista Subang.</p>	
6	<p>YongLiang Deng, Ying Zhang, Zhenmin Yuan, Rita Yi Man Li, and Tiantian Gu, 2023</p>	<p>Analyzing Subway Operation Accidents Causation: Apriori Algorithm and Network Approaches</p>	<p>Penelitian ini bertujuan untuk mengusulkan kerangka analitis terintegrasi untuk mengidentifikasi dan menganalisis hubungan antara kecelakaan dan penyebabnya dalam operasi kereta bawah tanah. Menerapkan algoritma apriori untuk menambang asosiasi potensial antara faktor penyebab kecelakaan, menggunakan teori jaringan kompleks untuk membangun jaringan kausalitas kecelakaan operasi kereta bawah tanah (SOACN), dan</p>	<p>Kekurangan pada penelitian ini ialah adanya keterbatasan dalam penentuan bobot simpul dan tepi serta perlunya penelitian lebih lanjut terkait probabilitas risiko dan langkah pengendalian dalam situasi sumber daya terbatas.</p>

			<p>menganalisis fitur topologi dan kerentanan SOACN. Hasil penelitian ini adalah mengembangkan model jaringan SOACN yang mengintegrasikan kecelakaan dan penyebab dalam operasi kereta bawah tanah dengan pendekatan data mining. Temuan menunjukkan perbedaan karakteristik antara kecelakaan dan penyebab, serta pentingnya faktor mekanikal dalam pengendalian resiko.</p>	
7	Fitri Rahmawati, Nita Merlina, 2018	Metode Data Mining Terhadap Data Penjualan <i>Spare part</i> Mesin <i>Fotocopy</i> Menggunakan Algoritma Apriori	<p>Penelitian ini bertujuan untuk mengidentifikasi produk <i>spare part</i> yang banyak terjual secara bersamaan dan membentuk aturan asosiasi pada data penjualan. Metode yang digunakan adalah algoritma apriori. Hasil</p>	<p>Kekurangan jurnal ini adalah jurnal ini tidak memberikan informasi yang detail mengenai proses pengumpulan data dan analisis yang dilakukan.</p>

			dari penelitian jurnal ini ditemukan bahwa produk <i>Toner</i> dan <i>Drum Dc</i> adalah produk yang paling sering terjual secara bersamaan.	
--	--	--	--	--

### 2.1.1 Literatur (Sopiyon Ependi, Muhamad Akbar., 2021)

Penelitian yang dilakukan oleh Sopiyon Ependi, Muhamad Akbar (2021) dari Fakultas Ilmu Komputer, Universitas Bina Darma dengan judul Implementasi Data Mining Pada Penjualan Produk Dengan Menggunakan Algoritma Apriori. Dalam penelitian yang dilakukan oleh penulis, penulis mengangkat masalah yaitu CV Berkat Motor mengalami kesulitan untuk mengetahui seberapa banyak suku cadang dan jasa yang telah terjadi dalam satu transaksi, sehingga pemilik masih manual/spekulasi dalam pembelian suku cadang. Oleh karena itu CV Berkat Motor bertujuan untuk mendapatkan informasi mengenai pola transaksi dan mengetahui jasa dan penjualan produk suku cadang apa saja yang sering muncul dari CV Berkat Motor sehingga bisa membantu pimpinan dalam pengambilan keputusan untuk melakukan pembelian suku cadang yang perlu diiprioritaskan. Penelitian ini menggunakan metode Algoritma apriori. Teknik pengumpulan data pada jurnal ini meliputi dokumentasi, wawancara, studi kepustakaan. Teknik Analisa data yang digunakan pada penelitian ini adalah menggunakan tahapan *knowledge Discovery in Databases (KDD)* dan tahapan ini meliputi *Data Selection, Pre-Processing* atau *Cleaning*, Transformasi Data, Data Mining, *Interpratation* atau

*Evaluation.* Hasil yang diperoleh pada penelitian tersebut adalah manajer dapat mengambil keputusan yang dilakukan untuk menentukan strategi penjualan karena keputusan yang diambil akan menentukan peningkatan penjualan barang yang diproduksi oleh perusahaan menggunakan perangkat lunak R studio versi 3.4.

### **2.1.2 Literatur (Mukhlis Ramadhan, Juniar Hutagalung, Muhammad Dahria, Iskandar Zulkarnain., 2023)**

Penelitian yang dilakukan oleh Mukhlis Ramadhan, Juniar Hutagalung, Muhammad Dahria, Iskandar Zulkarnain (2023) dari STMIK Triguna Dharma dengan judul *Prediksi Penjualan Spare Part Mobil Daihatsu Menggunakan Algoritma Apriori*. Dalam penelitian yang dilakukan oleh peneliti, peneliti mengangkat masalah yaitu penurunan penjualan spare part mobil Daihatsu dan ketidaktepatan dalam menentukan strategi promosi yang diberikan kepada pelanggan. Banyaknya data transaksi yang digunakan sebagai acuan dalam menjual produk dengan harga modal yang hanya mendapatkan keuntungan kecil menjadi salah satu kendala dalam menganalisa penjualan dan strategi yang diberikan kepada pelanggan. Oleh karena itu, perusahaan membutuhkan sistem yang dapat mengolah data informasi yang lebih cepat dan tepat dalam meningkatkan penjualan spare part mobil Daihatsu. Hal ini bertujuan untuk mempermudah analisis terhadap daya saing produk spare part mobil Daihatsu yang paling laku terjual secara bersamaan, serta memberikan rekomendasi kepada pengambil keputusan untuk meningkatkan pemasaran dan promosi produk spare part mobil Daihatsu yang lebih baik. Metode analisis asosiasi dengan menggunakan algoritma apriori, metode ini digunakan untuk

menganalisis pola penjualan spare part mobil dan memprediksi pola penjualan yang paling laku terjual secara bersamaan. Hasil perhitungan dengan menerapkan algoritma apriori menunjukkan bahwa jika konsumen membeli lower arm daihatsu dan ban mobil maka nilai support = 23,33 dan nilai confidence = 77,78 dan jika konsumen membeli lower arm daihatsu dan filter AC maka nilai support = 26,67 dan nilai confidence = 72,72.

### **2.1.3 Literatur (Silvi Sintia, Poningsih, Ilham Syahputra Saragih, Anjar Wanto, Irfan Sudahri Damanik., 2019)**

Penelitian yang dilakukan Silvi Sintia, Poningsih, Ilham Syahputra Saragih, Anjar Wanto, Irfan Sudahri Damanik (2019) dari STIKOM Tunas Bangsa Indonesia dengan judul Penerapan Algoritma Apriori Dalam Memprediksi Hasil Penjualan Sparepart PC (Studi Kasus : Toko Sentra Computer). Dalam penelitian yang dilakukan oleh peneliti, peneliti mengangkat masalah Toko Sentra Computer sering mengalami kehabisan stok sparepart PC yang harganya diatas Rp. 450.000 karena stok yang dibeli dengan jumlah yang sedikit agar modal bias dimanfaatkan ke sparepart PC yang lebih murah. Hal ini menyebabkan kesulitan bagi pemilik Toko dalam memenuhi permintaan konsumen. Untuk mengetahui data penjualan sparepart PC pada Toko Sentra Computer yang paling banyak dibeli oleh konsumen. Dengan menggunakan algoritma apriori, penelitian ini bertujuan untuk membantu pemilik Toko dalam memprediksi hasil kombinasi penjualan sparepart PC dan meningkatkan ketersediaan stok barang. Algoritma apriori adalah salah satu algoritma pada bidang data mining untuk penggalian asosiasi atau yang lebih dikenal dengan istilah *association rule mining*, tahap ini mencari kombinasi item yang

memenuhi syarat minimum dari nilai *support* dalam basis data. Algoritma apriori digunakan untuk menggali aturan asosiasi atau *association rule mining* dari kumpulan data penjualan. Hasil akhir *confidence* akan dibandingkan dengan hasil akhir perhitungan menggunakan *software rapidminer* dengan nilai minimum *support* dan *confidence* yang sama, berdasarkan hasil perhitungan pada penelitian tersebut didapat nilai *confidence ram* adalah yang paling tinggi dan perbandingan hitungan manual dengan menggunakan *software rapidminer* memberikan hasil yang sama. Oleh karena itu pimpinan Toko Sentra Computer dapat menyediakan stok sparepart PC yang paling laku agar tercukupi dengan minat konsumen.

#### **2.1.4 Literatur (P. Aiswarya dkk., 2020)**

Penelitian ini berjudul “Detection and Prediction of Frequent Diseases in India through Association Technique using Apriori Algorithm and Random Forest Regression” perubahan profil penyakit di india dari penyakit menular menjadi penyakit tidak menular, memerlukan pengambilan keputusan yang efektif dalam pengendalian penyakit. Penelitian ini bertujuan untuk mengidentifikasi dan memprediksi penyakit yang sering terjadi di india menggunakan teknik data mining. Penelitian ini menggunakan metode data mining algoritma apriori untuk mengidentifikasi dan memprediksi penyakit yang sering terjadi di india. Hasil dari penelitian tersebut adalah penggunaan teknik data mining seperti algoritma apriori dapat membantu dalam mengidentifikasi dan memprediksi penyakit yang sering terjadi di india, penelitian tersebut dapat digunakan untuk pengambilan keputusan yang lebih efektif dalam pengendalian penyakit dan alokasi sumber daya kesehatan.

Penelitian ini memberikan kontribusi penting dalam meningkatkan perawatan kesehatan dan fasilitas bagi pasien di india.

### **2.1.5 Literatur (Adi Sucipto, Hadapiningradja Kusumodestoni, 2022)**

Penelitian yang dilakukan Adi Sucipto, Hadapiningradja Kusumodestoni (2022) dari Program Studi Teknik Informatika, Universitas Islam Nahdlatul Ulama Jepara dengan judul Implementasi Data Mining Algoritma Apriori Terhadap Penjualan Spare Part Motor di Yamaha Arista Subang. Proses pengolahan data penjualan Spare Part yang dilakukan oleh Yamaha Arista Subang masih menggunakan cara yang sederhana dan tidak efektif. Hal ini menyebabkan kesulitan dalam melihat pola asosiasi antara penjualan spare part satu dengan spare part yang lainnya. Contih kasus yang sering terjadi ialah stok spare part kosong yang merupakan spare part *fast moving* yang sering dicari oleh konsumen, sehingga mengakibatkan turunnya tingkat kepercayaan pelanggan terhadap bengkel. Oleh karena itu, diperlukan sistem pengolah database melalui aplikasi data mining yang dapat membantu bengkel dalam mengambil keputusan secara cepat dan tepat. Membangun sistem informasi menggunakan metode algoritma apriori dalam menentukan pola kombinasi barang yang sering terjual pada penjualan spare part dibengkel Yamaha Arista Subang. Sistem ini diharapkan dapat membantu bengkel dalam mengambil keputusan yang cepat dan tepat dalam mengatur stok spare part serta meningkatkan strategi pemasaran dan pemodelan penjualan spare part motor. Algoritma apriori digunakan untuk menganalisi pola asosiasi antara penjualan spare part satu dengan spare part lainnya berdasarkan data transaksi penjualan yang telah dikumpulkan. Sistem dapat menghasilkan aturan asosiasi

menggunakan 206 data penjualan spare part motor dengan batasan *minimum support* 15% dan *minimum confidence* 30%, sehingga menghasilkan 6 pola aturan asosiasi untuk penjualan spare part motor pada bulan juni 2020 di Yamaha Arista Subang.

#### **2.1.6 Literatur (Yongliang Deng, Ying Zhang, Zhenmin Yuan, Rita Yi Man Li, and Tiantian Gu.,2023)**

Penelitian ini berjudul “Analyzing Subway Operation Accidents Causations: Apriori Algorithm and Network Approaches” berfokus pada kompleksitas dan dinamika manajemen keselamatan oprasi kereta bawah tanah. Masalah utama yang dihadapi adalah memahami penyebab kecelakaan dengan lebih baik. Oleh karena itu, penelitian ini menggabungkan algoritma apriori dengan teori jaringan kompleks untuk menganalisis penyebab kecelakaan. Tujuannya adalah mengembangkan kerangka analitis yang terintegrasi untuk mengidentifikasi dan menganalisis hubungan antara kecelakaan dan penyebabnya. Dengan mengklasifikasikan kecelakaan menjadi 13 jenis, penelitian ini menggali faktor penyebab melalui analisis kasus dan pengalaman ahli. Melalui penerapan algoritma apriori, penelitian ini menemukan potensi asosiasi antara faktor penyebab dan jenis kecelakaan. Dengan membangun jaringan penyebab kecelakaan operasi kereta bawah tanah (SOACN) menggunakan teori jaringan kompleks, penelitian ini mengidentifikasi fitur topologi dan kerentanan SOACN. Hasil penelitian memberikan pandangan yang dalam, menunjukkan bahwa beberapa faktor penyebab memiliki dampak lebih besar pada keselamatan operasi, dan rekomendasi yang dihasilkan

menekankan pada peningkatan pengawasan, pelatihan karyawan, dan sistem manajemen keselamatan.

### **2.1.7 Literatur (Fitri Rahmawati, Nita Merlina.,2018)**

Penelitian ini berjudul “Metode Data Mining Terhadap Data Penjualan *Spare Part* Mesin *Fotocopy* Menggunakan Algoritma Apriori” penjualan produk *spare part* mesin fotocopy semakin diperlukan, namun data penjualan tersebut tidak dimanfaatkan dengan baik oleh perusahaan. Data penjualan hanya berfungsi sebagai arsip perusahaan dan tidak digunakan untuk pengembangan strategi pemasaran. Oleh karena itu, diperlukan metode yang dapat mengidentifikasi produk *spare part* mesin fotocopy yang banyak terjual secara bersamaan untuk meningkatkan tingkat penjualan. Dengan mengaplikasikan metode data mining menggunakan algoritma apriori pada data penjualan *spare part* mesin fotocopy ini bertujuan untuk mengidentifikasi produk *spare part* yang banyak terjual secara bersamaan dan membentuk aturan asosiasi pada data penjualan. Hasil penelitian dalam jurnal ini menunjukkan bahwa produk *Toner* dan *Drum Dc* adalah produk yang sering terjual secara bersamaan. Hal ini menunjukkan adanya hubungan atau asosiasi antar kedua produk tersebut. Jurnal ini juga menunjukkan bahwa algoritma apriori efektif dalam menemukan hubungan antara produk dan mengidentifikasi produk terlaris.

## **1.2 Penjualan**

Penjualan adalah suatu upaya yang terintegrasi dalam merancang strategi-strategi yang bertujuan untuk memenuhi kebutuhan dan keinginan pembeli dengan tujuan menghasilkan keuntungan (Kotler & Kevin Lane Keller, 2016). Penjualan merupakan aspek utama dalam kelangsungan hidup suatu

perusahaan, karena melalui penjualan inilah perusahaan memperoleh pendapatan dan menarik konsumen untuk mengenal produk yang ditawarkan sehingga dapat menilai hasil produk tersebut. Penjualan adalah jumlah uang yang dibayarkan oleh pembeli untuk barang dagang yang diberikan, yang kemudian menjadi pendapatan bagi perusahaan yang bersangkutan (Indarwati dkk., 2018).

### **1.3 Prediksi**

Prediksi merupakan suatu proses yang dilakukan secara terencana untuk mengestimasi kemungkinan terjadinya suatu peristiwa di masa depan berdasarkan data masa lalu dan saat ini tersedia. Tujuan utamanya adalah mengurangi tingkat ketidakpastian dengan meminimalkan kesalahan dalam memprediksi perbedaan antara hasil prediksi dengan kenyataan yang terjadi. Prediksi tidak selalu memberikan jawaban pasti mengenai peristiwa yang akan datang, tetapi berupaya untuk memberikan perkiraan yang seakurat mungkin (Orpa dkk., 2019).

### ***1.4 Spare part***

*Spare part* atau suku cadang adalah sebuah produk yang terdiri dari berbagai komponen yang bekerja bersama untuk melaksanakan fungsi tertentu. Setiap *spare part* memiliki fungsi uniknya sendiri dan dapat berhubungan atau terpisah dari komponen lainnya. Secara umum, *spare part* dapat diklasifikasikan menjadi dua jenis (Welim dkk., 2015), yaitu:

1. *Spare part* baru adalah komponen yang masih dalam kondisi baru dan belum pernah digunakan, kecuali saat pengujian.

2. *Spare part* bekas atau bekas pakai adalah komponen yang telah digunakan selama periode atau jangka waktu tertentu dan dapat dikelompokkan dalam dua kondisi yaitu layak pakai dan tidak layak pakai.

### **1.5 Bengkel**

Bengkel merupakan lokasi yang digunakan untuk melakukan perbaikan, pemeliharaan, dan perancangan mesin, dengan berbagai peralatan konstruksi dan komponen mesin yang tersedia di dalamnya. Pengetahuan dan keterampilan terkait bengkel juga dapat disebut sebagai keahlian dalam perbengkelan (Novianto & Santoso, 2018).

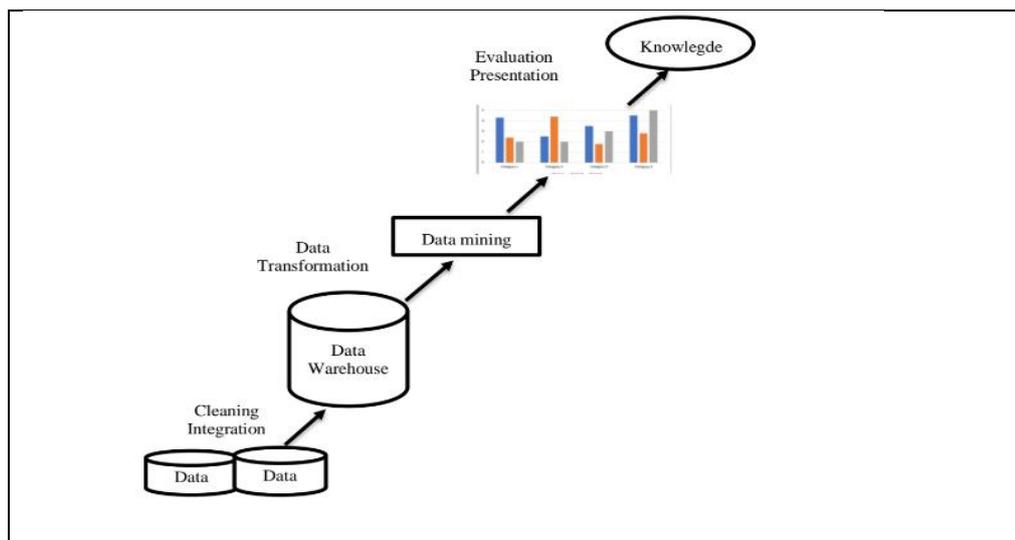
### **1.6 Data Mining**

Data mining merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data (Larose, 2005). Data mining juga merupakan suatu proses yang dalam penerapannya menggunakan teknik statistik, matematika, kecerdasan buatan, *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Efraim Turban dkk., 2005). Data mining ialah salah satu bidang ilmu yang perkembangannya sangat pesat, perkembangan tersebut disebabkan oleh beberapa faktor (Bulolo, 2020), antara lain:

1. Pentingnya akan kesadaran data semakin tinggi
2. Pemanfaatan hasil dari pengolahan data dalam berbagai bidang salah satu contohnya di bidang bisnis
3. Berkembangnya kumpulan data yang begitu cepat

4. Meningkatnya akses internet baik melalui web ataupun *smartphone*
5. *Hardware* dan *software* yang khususnya memiliki keterkaitan dengan data mining semakin berkembang
6. Komputasi komputer yang berkembang begitu cepat
7. Perkembangan media penyimpanan yang semakin besar dan harga yang relatif terjangkau

Tahapan proses dalam data mining yang merujuk pada proses *Knowledge Discovery in Database (KDD)* seperti yang dapat dilihat pada Gambar 2.1 dapat dijelaskan sebagai berikut:



Gambar 2. 1 Proses KDD (Builolo, 2020)

Tahapan *KDD* dapat diuraikan menjadi beberapa tahapan, (Builolo, 2020) yaitu sebagai berikut:

1. Data

Pada proses *KDD* pertama kali yang harus dipersiapkan adalah data. Data yang dimaksud adalah data yang sudah terpisah dengan data operasional.

## 2. *Selection*

*Selection* atau pemilihan pada data meliputi pembulatan kumpulan data target, menentukan variabel, pemilihan sampel data dan penyimpanan data di dalam sebuah berkas.

## 3. *Pre-Processing /Cleaning*

Proses pembersihan atau *cleaning* mencakup pembuangan duplikasi data, memperbaiki data yang inkonsisten, dan memperbaiki kesalahan data. Pada tahapan ini juga dapat dilakukan proses memperkaya data dengan menambah informasi yang relevan atau dengan kata lain *enrichment*.

## 4. *Transformation*

Data yang dipersiapkan untuk digunakan dalam proses *KDD*, diubah terlebih dahulu menyesuaikan dengan metode atau algoritma yang akan digunakan.

## 5. *Data mining*

Proses penggalian serta pencarian informasi yang bermanfaat menggunakan metode atau algoritma yang sesuai dengan informasi yang dicari.

## 6. *Interpretation/Evaluation*

Informasi yang dihasilkan dari tahap data mining akan disajikan atau ditampilkan dalam bentuk yang lebih mudah dipahami seperti bentuk grafik, pohon keputusan atau dalam bentuk *rule*. Lalu hasil dari data mining akan diperiksa apakah bertentangan atau tidak dengan hipotesa ataupun fakta yang sebelumnya.

## 7. Knowledge

Tujuan dari proses *KDD* adalah ingin memperoleh informasi yang mudah dipahami dan bermanfaat. Informasi yang dihasilkan akan diimplementasikan sesuai kebutuhan, manfaat, dan juga kegunaan informasi tersebut.

### 1.7 Algoritma Apriori

Algoritma apriori merupakan teknik data mining untuk menemukan aturan asosiasi antara suatu kombinasi item (Sitanggang, 2023). Penting tidaknya suatu asosiasi dapat diketahui dengan dua tolak ukur, yaitu : *support* dan *confidence*. *Support* (nilai penunjang) adalah presentase kombinasi *item* dalam database, sedangkan *confidence* (nilai kepastian) adalah kuatnya hubungan antar-*item* dalam aturan asosiasi (Umami, 2016).

#### 1. Analisis Pola Frekuensi Tinggi dengan Algoritma Apriori

Tahap ini mencari kombinasi *item* yang memenuhi syarat minimum dari nilai *support* dalam basis data. Nilai *support* sebuah item diperoleh dengan menggunakan rumus berikut :

$$\text{Support A} = \frac{\sum \text{Transaksi mengandung A}}{\sum \text{Transaksi}} \times 100\% \quad (2.1)$$

Nilai *support* dari 2 *item* diperoleh dengan rumus :

$$\text{support (A, B)} = P(A \cap B) \quad (2.2)$$

$$\text{Support A, B} = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi}} \times 100\%$$

## 2. Pembentukan Aturan Asosiasi

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif  $A \rightarrow B$ . Nilai dari aturan *confidence*  $A \rightarrow B$  diperoleh dengan rumus :

$$support = \frac{\sum \text{Transaksi mengandung } A \text{ dan } B}{\sum \text{Transaksi mengandung } A} \quad (2.3)$$

Untuk menentukan aturan asosiasi yang akan dipilih maka harus diurutkan berdasarkan  $Support \times Confidence$ . Aturan diambil sebanyak  $n$  aturan yang memiliki hasil terbesar (Umami, 2016).

### **1.8 Association Rule**

Analisis asosiasi atau penambangan aturan asosiasi adalah metode data mining yang digunakan untuk mengidentifikasi hubungan antara berbagai kombinasi item. Analisis asosiasi juga sering disebut sebagai fondasi bagi beberapa teknik data mining. Sebab bermula dari penelitian tentang data base transaksi konsumen guna mengidentifikasi kecenderungan pembelian produk bersama, maka *association rule* sering juga disebut *market basket analysis* (Sitanggang, 2023). Aturan asosiasi bertujuan untuk memberikan informasi dalam format hubungan “jika-maka” atau “*if-then*”. Pentingnya sebuah aturan asosiasi dapat diidentifikasi melalui dua parameter, yakni *support* (nilai penunjang) yang merupakan presentase kombinasi item dalam basis data dan *confidence* (nilai kepastian) yaitu tingkat keterikatan antar item dalam aturan asosiatif (Dian Wirdasari & Ahmad Calam, 2011).

### **1.9 Google Colaboratory**

*Google Colaboratory* atau *google colab* merupakan platform yang memungkinkan pengguna untuk menulis dan menjalankan kode *python* melalui peramban web. Platform ini dibuat oleh *google* dan diperuntukkan khusus dalam pengembangan *machine learning*, *deep learning*, dan analisis data. Kelebihan utama dari *google colab* adalah tidak memerlukan pengaturan khusus, memberikan akses gratis ke sumber daya pemrosesan umum, serta sifatnya yang mudah untuk dibagikan. Dengan menggunakan *google colaboratory*, para pengembang dapat menggabungkan kode yang dapat dieksekusi dengan teks berformat dalam satu dokumen tunggal, termasuk gambar, HTML, dan LaTeX. Tambahan pula, *google colab* memungkinkan kolaborasi online antara pengguna melalui berbagai fitur pengkodean (Rohim & Zuliarso, 2022).

### **1.10 Python**

*Python* adalah bahasa pemrograman yang diciptakan oleh Guido van Rossum, dengan ciri khas sebagai bahasa berorientasi objek dan tingkat tinggi. Keunggulan utama terletak pada kemampuan mendukung berbagai paradigma pemrograman seperti berorientasi objek, imperatif, dan fungsional. Selain itu, *python* juga terkenal karena sintaksis kode yang sangat jelas dan mudah dimengerti. Bahasa ini dilengkapi dengan beragam pustaka yang luas dan komprehensif. Kelebihan lainnya adalah kemampuannya sebagai bahasa pemrograman dinamis dengan pengelolaan memori otomatis. Meskipun umumnya digunakan untuk scripting, *python* memiliki aplikasi yang lebih luas dan mencakup berbagai bidang seperti pengembangan aplikasi, analisis data,

kecerdasan buatan, dan sebagainya. Karena berbagai keunggulan ini, *python* menjadi salah satu bahasa pemrograman yang sangat populer dan banyak digunakan dalam berbagai konteks (Nurjanah & Insanudin, 2016).