

BAB II

TINJAUAN PUSTAKA

2.1 Teori dasar

Teori dasar dalam penelitian ini akan membahas teori-teori yang berhubungan dengan topik penelitian yang diusulkan, pembahasan teori ini menjadi penting sebagai instrumen untuk meningkatkan validitas dari penelitian yang dilaksanakan. Dalam pembahasannya teori-teori yang dibahas terdiri dari pengertian dan istilah umum teori yang digunakan, adapun teori yang akan dibahas dalam penelitian ini terdiri dari: (1) *Knowledge Discovery in Database* (KDD), (2) *Data Mining*, (3) *Algoritma Association Rules*, (4). *Algoritma Apriori*, (5) Pola Pembelian Konsumen, (6) *Google Colaboratory*, dan (7) *Database*.

2.1.1 *Knowledge Discovery in Database* (KDD)

Seiring dengan meningkatnya volume dan kompleksitas data yang dimiliki oleh berbagai entitas bisnis, kebutuhan untuk mengekstraksi informasi bernilai dari data tersebut menjadi semakin penting. Salah satu pendekatan yang digunakan untuk menjawab tantangan ini adalah *Knowledge Discovery in Databases* (KDD), yaitu proses sistematis untuk mengekstrak pengetahuan tersembunyi dari kumpulan data berskala besar (Anggrawan et al., 2021).

Knowledge Discovery in Databases (KDD) terdiri dari beberapa tahapan penting seperti seleksi data, praproses data, transformasi, proses *Data Mining*, dan evaluasi hasil untuk memperoleh pola atau pengetahuan yang dapat diinterpretasikan. Tujuan dari implementasi KDD adalah untuk membantu

organisasi dalam memperoleh informasi bermakna dari data yang tersimpan secara digital melalui pendekatan eksploratif dan terstruktur. Salah satu tahapan inti dalam KDD adalah *Data Mining*, yakni proses untuk menemukan pola, keterkaitan, atau asosiasi yang signifikan dalam kumpulan data (Merliani et al., 2022).

Data Mining dalam bisnis sangat berperan dalam memahami perilaku pelanggan, mengidentifikasi kebiasaan pembelian, dan memformulasikan strategi pemasaran berbasis data (Ndoya & Putri, 2024). Hal ini selaras dengan pandangan lain yang menekankan bahwa penerapan KDD melalui algoritma seperti *Apriori* dapat digunakan dalam praktik *market basket analysis* untuk mendeteksi kombinasi item yang sering muncul dalam satu transaksi, yang pada akhirnya membantu dalam optimalisasi manajemen persediaan dan pengambilan keputusan operasional.

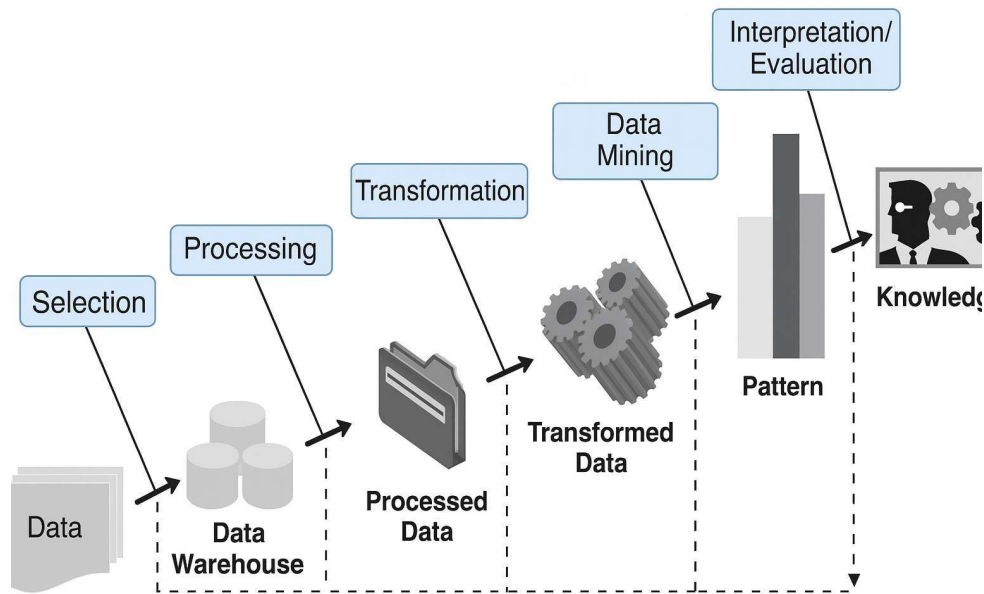
Tahapan-tahapan yang terdapat dalam KDD terdiri dari 5 tahapan utama, tahapan-tahapan yang terdapat dalam KDD adalah sebagai berikut (Alghifari & Juardi, 2021):

1. *Data Selection*, tahapan pemilihan data (*data selection*) dilakukan dengan tujuan memilih *subset* data yang relevan dan *representatif* untuk dilakukan analisis lebih lanjut, hal ini diperlukan untuk mengurangi *noise* yang dikandung dalam data. Pemilihan data yang baik akan menghasilkan proses pengolahan dan luaran yang lebih baik.
2. *Pre-processing*, tahapan pra-proses dalam KDD merupakan langkah-langkah yang dilakukan sebelum melanjutkan ke tahap penambangan data (*Data Mining*). Tahapan ini sangat penting untuk mempersiapkan data yang akan digunakan dalam analisis. Tahapan *pra-proses* biasanya terdiri dari beberapa

tahapan penting seperti pembersihan data, dan melakukan perbaikan dalam data.

3. *Transformation*, pada tahap ini akan dilakukan proses pencarian fitur-fitur yang bermanfaat untuk mewakili data tergantung pada tujuan yang ingin dicapai, langkah transformasi data yang penting setelah pemilihan data yang relevan. Transformasi ini mengubah data yang telah dipilih agar sesuai dengan persyaratan yang diperlukan dalam proses penambangan data. Proses ini memiliki unsur kreatif dan sangat bergantung pada jenis informasi atau pola yang akan diidentifikasi dalam basis data
4. *Data Mining*, pada tahapan ini dilakukan berbagai pemilihan mulai dari metode *Data Mining* yang akan digunakan, seperti *klasifikasi*, *regresi*, *clustering*, dll. Serta melakukan pemilihan algoritma untuk melakukan pencarian. Proses *Data Mining* adalah upaya untuk mencari pola tertentu untuk informasi menarik dalam data terpilih dengan menggunakan berbagai teknik, metode atau algoritma yang sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.
5. *Interpretation/Evaluation*, tahapan ini menampilkan hasil penjumlahan pola (*pattern*) yang dihasilkan dari proses *Data Mining*, tahapan ini juga akan melakukan evaluasi dari pola atau informasi yang didapat sesuai atau bertentangan dengan fakta/hipotesa yang ada sebelumnya

Adapun proses *Knowledge Discovery in Database* hasil penjabaran di atas dapat dilihat pada gambar di bawah ini:



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

Sumber: (Walhaidayat et al, 2021)

2.1.2 Data Mining

Data Mining merupakan inti dari proses *Knowledge Discovery in Databases* (KDD) yang berfokus pada penggalian informasi tersembunyi dari data dalam jumlah besar melalui pendekatan statistik, kecerdasan buatan, dan algoritma pembelajaran mesin. *Data Mining* adalah proses untuk menemukan pola atau hubungan baru yang sebelumnya tidak diketahui dalam kumpulan data besar, dengan tujuan menghasilkan pengetahuan yang dapat digunakan secara praktis (Anggrawan et al., 2021).

Proses ini bukan sekadar pencarian data, melainkan sebuah pendekatan sistematis yang melibatkan beberapa tahap penting, seperti transformasi data, pemilihan atribut, hingga penerapan algoritma analitik untuk mengidentifikasi

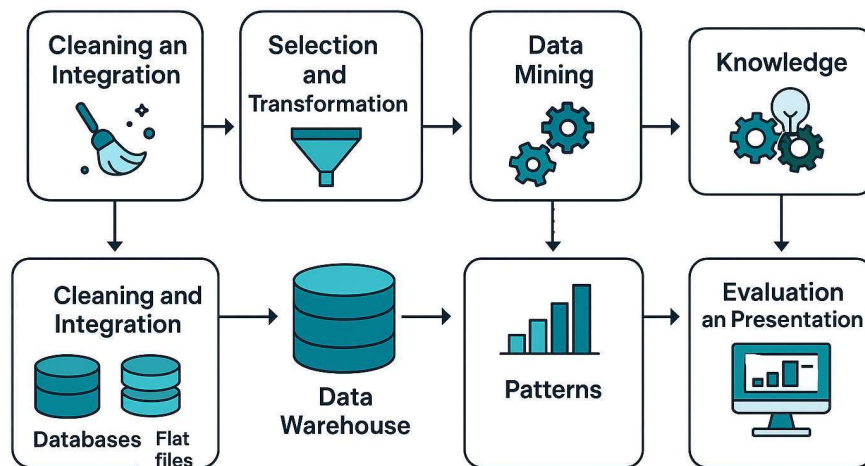
keterkaitan atau pola bermakna. menyebutkan bahwa *Data Mining* merupakan metode eksplorasi dan analisis sistematis terhadap data yang tersimpan, yang memungkinkan pengambilan keputusan strategis berbasis informasi historis (Merliani et al., 2022).

Data Mining dapat digunakan dalam konteks bisnis untuk memahami perilaku konsumen, memetakan pola pembelian, dan mendukung penyusunan strategi pemasaran berbasis data yang lebih efisien dan relevan dengan kebutuhan pelanggan (Ndoya & Putri., 2024).

Pada *Knowledge Discovery in Databases* (KDD), proses *Data Mining* memiliki beragam jenis pendekatan yang dapat disesuaikan dengan kebutuhan analisis data. Secara umum, fungsi dari *Data Mining* mencakup beberapa tugas utama, seperti deskripsi, prediksi, klasifikasi, pengelompokan, estimasi, dan asosiasi. Pendekatan deskriptif digunakan untuk menjelaskan karakteristik data dan pola-pola yang dapat dipahami secara intuitif oleh manusia. Estimasi dan prediksi berfokus pada upaya memperkirakan nilai suatu atribut atau kejadian berdasarkan pola historis yang ditemukan dalam data. Klasifikasi bertujuan untuk mengelompokkan data ke dalam kategori yang telah ditentukan, sedangkan pengelompokan (*clustering*) mengidentifikasi struktur tersembunyi dalam data tanpa label yang sudah ada. Sementara itu, metode asosiasi digunakan untuk menemukan hubungan atau keterkaitan antar item yang muncul bersamaan, yang sangat bermanfaat dalam analisis keranjang belanja (*market basket analysis*) (Lewis et al., 2021).

Proses KDD mencakup beberapa tahapan utama yang dimulai dari pembersihan data (*data cleaning*), yaitu data yang tidak konsisten, duplikat, atau *error* dibersihkan agar analisis menjadi akurat. Tahap berikutnya adalah integrasi data (*data integration*), yaitu menggabungkan berbagai sumber data ke dalam satu basis data yang seragam. Setelah itu, dilakukan seleksi data (*data selection*) untuk memilih atribut atau data yang relevan dengan tujuan analisis. Data yang telah diseleksi kemudian diproses pada tahap transformasi (*data transformation*), agar sesuai dengan format atau struktur yang dibutuhkan oleh metode *Data Mining* yang akan digunakan (Merliani et al., 2022).

Tahap inti dari proses ini adalah *Data Mining* itu sendiri, yaitu penerapan algoritma tertentu untuk menemukan pola atau hubungan penting dalam data. Setelah pola ditemukan, tahap evaluasi pola (*pattern evaluation*) dilakukan untuk menilai sejauh mana pola tersebut signifikan dan dapat digunakan dalam konteks pengambilan keputusan. Hasil dari analisis tersebut disajikan melalui tahap presentasi pengetahuan (*knowledge presentation*), yang bertujuan menyampaikan temuan secara jelas dan informatif kepada pengguna, sering kali melalui visualisasi atau laporan yang mudah dipahami. Seluruh tahapan ini menjadikan KDD sebagai pendekatan sistematis dan berulang untuk menggali pengetahuan tersembunyi dari kumpulan data yang kompleks dan besar (Prasetya et al., 2021).



Gambar 2. 2 Tahapan *Data Mining*
Sumber: (Walhaidayat et al, 2021)

Berdasarkan gambar di atas, maka tahapan-tahapan *data mining* dapat dijelaskan sebagai berikut:

1. *Cleaning and integration* bertujuan untuk menghilangkan data yang tidak konsisten, hilang, atau duplikasi agar kualitas terjamin yang dilakukan dengan cara membersihkan *noise/error* di dalam data.
2. *Selection and Transformation* bertujuan untuk memilih data yang relevan dengan tujuan analisis. Proses ini meliputi pemilihan subset data penting, normalisasi, agregasi, serta transformasi data agar lebih sesuai dengan kebutuhan analisis.
3. *Data Warehouse* bertujuan untuk menyimpan data hasil integrasi dalam bentuk yang terstruktur. Data warehouse menjadi pusat

penyimpanan yang memudahkan akses, manajemen, dan analisis data secara konsisten.

4. *Data mining* bertujuan untuk menggali pola atau hubungan yang tersembunyi di dalam data dengan menggunakan algoritma tertentu, seperti *association rules*, *clustering*, *classification*, atau regresi.
5. *Pattern* bertujuan untuk menampilkan pola atau aturan yang ditemukan dari hasil data mining. Pola ini bisa berupa hubungan antar item, kelompok pelanggan, atau tren tertentu yang muncul dari data.
6. *Evaluation and Presentation* bertujuan untuk mengevaluasi pola yang ditemukan agar dapat dipastikan validitas, kegunaan, dan signifikansinya. Hasil evaluasi kemudian dipresentasikan dalam bentuk tabel, grafik, maupun visualisasi lain agar mudah dipahami.
7. *Knowledge* bertujuan untuk menghasilkan pengetahuan baru yang dapat digunakan dalam pengambilan keputusan strategis. Pengetahuan ini menjadi dasar untuk rekomendasi, perencanaan, dan tindakan yang lebih tepat sasaran.

2.1.3 Algoritma Association Rules

Association Rule merupakan salah satu metode dalam *Data Mining* yang bertujuan untuk menemukan hubungan tersembunyi antar item dalam suatu kumpulan data. Teknik ini sangat efektif digunakan dalam analisis perilaku konsumen, terutama dalam *market basket analysis*, yang menggunakan pola pembelian konsumen agar dapat dianalisis untuk mengidentifikasi keterkaitan antar

produk. Misalnya, dalam sebuah toko ritel, *Association Rule* dapat mengungkap bahwa pelanggan yang membeli roti cenderung juga membeli selai, sehingga informasi ini dapat dimanfaatkan dalam penataan produk atau strategi promosi. Pendekatan ini menyajikan aturan berbentuk implikasi logis *jika A maka B* yang disebut sebagai rule atau aturan asosiasi (Merliani et al., 2022).

Tiga ukuran penting dalam pembentukan *association rule* adalah *support*, *confidence*, dan *lift ratio*. *Support* menunjukkan seberapa sering kombinasi item tertentu muncul dalam dataset, *confidence* mengukur kekuatan hubungan antar item dalam sebuah aturan, sementara *lift ratio* menilai sejauh mana hubungan antar item tersebut lebih kuat dibandingkan jika terjadi secara acak. Ketiga metrik ini digunakan untuk mengevaluasi kualitas aturan yang terbentuk. Sebuah aturan dianggap kuat jika memiliki nilai *support* dan *confidence* yang tinggi serta *lift ratio* lebih besar dari 1, yang menunjukkan adanya korelasi positif antar item dalam aturan tersebut (Vidiya & Testiana, 2023).

Adapun persamaan-persamaan dari *support*, *confidence*, dan *lift ratio* adalah sebagai berikut:

$$\begin{aligned} \text{support}(A \rightarrow B) &= \frac{\text{Jumlah transaksi yang mengandung } A \text{ dan } B}{\text{Total transaksi}} \\ \text{confidence}(A \rightarrow B) &= \frac{\text{Jumlah transaksi yang mengandung } A \text{ dan } B}{\text{Jumlah transaksi yang mengandung } A} \\ \text{lift}(A \rightarrow B) &= \frac{\text{confidence}(A \rightarrow B)}{\text{support}(B)} \end{aligned}$$

Algoritma *Apriori* merupakan algoritma paling populer dalam proses pencarian *Association Rule*. Algoritma ini bekerja dengan cara membentuk

frequent itemset berdasarkan nilai minimum *support* yang ditentukan, lalu membangkitkan aturan berdasarkan nilai *minimum confidence* (Prasetya et al., 2021).

2.1.4 Algoritma *Apriori*

Algoritma *Apriori* merupakan salah satu algoritma paling mendasar dan banyak digunakan dalam *Association Rule mining*, yaitu teknik dalam *Data Mining* yang bertujuan untuk menemukan keterkaitan antar item dalam kumpulan data transaksi. Tujuan utama algoritma ini adalah untuk mengidentifikasi *frequent itemsets*, yaitu kombinasi item yang sering muncul bersama dalam satu transaksi, dan kemudian menghasilkan *Association Rules* yang memenuhi ambang batas minimum *support* dan *confidence* (Prasetya et al., 2021; Vidiya & Testiana, 2023).

Pendekatan *Apriori* sangat relevan dalam analisis *market basket analysis*, yang menggunakan pola pembelian konsumen dianalisis untuk membantu pengambilan keputusan strategis seperti penempatan produk, penawaran *bundling*, serta promosi yang ditargetkan. Selain itu, algoritma ini juga banyak diterapkan dalam bidang pendidikan, kesehatan, dan pemasaran untuk mengevaluasi pola tersembunyi dari data historis (Lewis et al., 2021).

Prinsip utama yang digunakan dalam algoritma *Apriori* adalah prinsip *downward closure*, yang menyatakan bahwa jika suatu kombinasi item (itemset) tidak memenuhi ambang batas minimum *support*, maka semua superset dari *itemset* tersebut juga tidak akan memenuhi syarat. Dengan memanfaatkan prinsip ini,

Apriori mampu mengeliminasi sejumlah besar kandidat *itemset* yang tidak relevan, sehingga meningkatkan efisiensi proses pencarian pola (Prasetya et al., 2021).

Proses algoritmik *Apriori* bersifat iteratif dan bertingkat. Proses dimulai dengan menghitung *support* untuk setiap item tunggal (*1-itemset*), kemudian dilanjutkan dengan pembentukan kandidat *2-itemset*, dan seterusnya. Setiap tahap pembentukan kandidat dilakukan dengan teknik *self-joining* dan diikuti dengan proses *pruning* terhadap kombinasi yang tidak memenuhi kriteria. Iterasi dihentikan ketika tidak ada lagi kandidat *itemset* yang bisa dibentuk (Merliani et al., 2022;).

Setelah *frequent itemsets* dihasilkan, algoritma *Apriori* melanjutkan dengan pembentukan aturan asosiasi. Pada tahap ini, sistem menghitung *confidence* dari setiap aturan yang mungkin dan hanya mempertahankan aturan yang memenuhi ambang batas minimum *confidence*. Dalam beberapa studi, metrik tambahan seperti *lift ratio* juga digunakan untuk mengukur kekuatan dan validitas hubungan antar item yang dihasilkan (Pratiwi et al., 2024).

Keunggulan algoritma *Apriori* terletak pada prosedurnya yang sistematis, mudah dipahami, serta fleksibel diterapkan pada berbagai domain. Namun, kelemahan utama dari pendekatan ini adalah kebutuhan untuk melakukan *scanning* berulang terhadap database, yang berpotensi meningkatkan beban komputasi, terutama pada dataset berukuran besar dengan jumlah item yang banyak. Untuk itu, beberapa pengembangan seperti *FP-Growth* diajukan sebagai alternatif yang lebih efisien.

2.1.5 Pola Pembelian Konsumen

Pola pembelian konsumen merupakan perilaku sistematis yang tercermin dari kebiasaan konsumen dalam membeli produk dalam satu atau lebih transaksi. Pola ini mencakup aspek seperti frekuensi pembelian, kombinasi produk yang dibeli secara bersamaan, hingga preferensi terhadap merek atau kategori produk tertentu. Analisis terhadap pola ini sangat penting dalam konteks pengambilan keputusan bisnis karena dapat mengungkap wawasan tersembunyi yang bermanfaat dalam mengatur strategi pemasaran, tata letak produk, manajemen stok, serta personalisasi promosi (Juliadi et al., 2024).

Pola pembelian konsumen dapat dikaji melalui pendekatan *Data Mining*, khususnya melalui metode *Association Rule Mining*. Salah satu tujuan utama dari pendekatan ini adalah untuk menemukan *itemset* yang sering dibeli secara bersamaan dalam satu transaksi. *Itemset* tersebut kemudian digunakan untuk membentuk aturan asosiasi yang merepresentasikan keterkaitan antar produk (Vidiya & Testiana, 2023).

Contohnya, jika sebuah aturan menyatakan bahwa “90% konsumen yang membeli oli mesin juga membeli busi dalam satu transaksi”, maka aturan ini sangat berguna bagi pemilik bisnis seperti untuk merancang *bundling* promosi, mengatur tata letak rak suku cadang, atau memberi rekomendasi layanan tambahan saat konsumen melakukan servis motor.

Pada AHHAS sendiri, pola pembelian konsumen umumnya melibatkan produk-produk seperti oli mesin, busi, filter udara, kampas rem, aki, ban, dan jasa servis berkala. Transaksi yang terjadi di AHHAS biasanya memiliki struktur tetap

yang memungkinkan analisis yang lebih akurat, misalnya, banyak pelanggan melakukan servis rutin setiap 2.000–4.000 km dan mengganti oli mesin bersamaan dengan item lain seperti busi atau kampas rem. Informasi seperti ini dapat digunakan untuk mendeteksi pola pembelian berulang, dan dengan algoritma seperti *Apriori*, pola-pola ini dapat diubah menjadi aturan *prediktif*.

Analisis pola pembelian juga dapat meningkatkan loyalitas pelanggan, dengan memberikan layanan berbasis histori pembelian. Misalnya, sistem bisa mengingatkan pelanggan untuk servis berikutnya berdasarkan pola pembelian sebelumnya, atau menawarkan diskon pada produk yang biasa dibeli bersamaan.

Hubungan antara pola pembelian konsumen dan teknik *Association Rule* sangat erat, karena *Association Rule* digunakan untuk mengekstraksi hubungan *implikatif* antar item dalam basis data transaksi. Penerapan algoritma seperti *Apriori*, pola-pola pembelian konsumen dapat dijabarkan menjadi aturan-aturan *prediktif* yang relevan. Sebagai contoh, jika pelanggan membeli *Oli MPX2*, maka kemungkinan besar mereka juga akan membeli filter udara. Demikian pula, apabila pelanggan melakukan *Servis CVT*, maka mereka cenderung mengganti *V-Belt* pada saat yang bersamaan. Penilaian terhadap kekuatan aturan asosiasi ini ditentukan oleh dua parameter utama, yaitu *support* dan *confidence*, dimana *Support* mengukur seberapa sering kombinasi item tertentu muncul dalam seluruh transaksi, sementara *confidence* menunjukkan *probabilitas* atau kemungkinan item B dibeli ketika item A telah dibeli. Kombinasi kedua nilai ini memungkinkan sistem untuk menyusun aturan yang tidak hanya valid secara statistik, tetapi juga bermanfaat untuk pengambilan keputusan strategis dalam konteks operasional dan pemasaran.

2.1.6 Google Colaboratory

Google Colaboratory (Colab) merupakan sebuah *platform* komputasi berbasis *cloud* yang dikembangkan oleh Google untuk mendukung aktivitas pemrograman dan eksplorasi data secara interaktif, terutama dengan menggunakan bahasa pemrograman Python. *Platform* ini memfasilitasi pengguna dalam mengembangkan, menjalankan, dan membagikan kode Python langsung melalui browser tanpa perlu mengatur lingkungan lokal, karena sudah terintegrasi dengan pustaka populer seperti *NumPy*, *pandas*, *scikit-learn*, *TensorFlow*, dan *Matplotlib*.

Google Colab sangat populer dalam bidang *Data Science*, *Machine Learning*, dan *Data Mining* karena menyediakan akses ke GPU secara gratis serta kemudahan kolaborasi melalui *Google Drive*. Pengguna dapat menyimpan dan mengakses *notebook* (dengan ekstensi *.ipynb*) secara langsung dari akun *Google Drive* mereka. Kelebihan ini menjadikan *Google Colab* sebagai pilihan ideal dalam pelatihan model, eksperimen algoritma, dan proses analitik data berskala besar, termasuk dalam konteks implementasi algoritma seperti *Apriori* dan *Association Rule Mining*.

Pada sejumlah penelitian, *Google Colab* telah digunakan sebagai sarana utama dalam eksperimen *Data Mining*. Misalnya memanfaatkan *Google Colab* untuk menguji model *Apriori* secara efisien menggunakan dataset transaksi, (Prasetya et al., 2021), sementara Putra et al. (2023) menggunakan platform ini untuk memvisualisasikan hasil proses mining dan mempercepat proses evaluasi pola dengan bantuan pustaka *Python* yang telah tersedia secara default. Kemudahan

dalam integrasi antara *notebook* dan lingkungan pemrosesan data membuat Colab menjadi alternatif ideal bagi peneliti maupun praktisi dalam bidang data analitik berbasis *cloud* (Prasetya et al., 2021; Vidiya & Testiana, 2023).



Gambar 2. 3 Google Colaboratory
Sumber: (Data Penelitian, 2025)

2.1.7 Database

Database atau basis data adalah kumpulan data yang terorganisir secara sistematis dan disimpan dalam media elektronik agar dapat diakses, dikelola, dan diperbarui dengan mudah. *Database* berperan penting sebagai fondasi utama dalam penyimpanan data transaksi, data pelanggan, maupun data operasional lainnya yang diperlukan untuk analisis dan pengambilan keputusan berbasis data (*data-driven decision making*).

Database dikelola oleh suatu perangkat lunak yang disebut *Database Management System* (DBMS). DBMS bertugas untuk memfasilitasi interaksi antara pengguna dan *database*, seperti saat melakukan proses *insert*, *update*, *delete*, maupun *query* terhadap *database*. DBMS juga berfungsi menjaga konsistensi, keamanan, dan integritas data dalam skala besar.

Database terdiri atas beberapa komponen utama seperti tabel untuk menyimpan data dalam bentuk baris dan kolom, *field/attribute* yaitu kolom dalam tabel yang mewakili jenis data tertentu, *record* yaitu baris dalam tabel yang merepresentasikan satu entitas data, *primary key* yaitu atribut unik yang digunakan untuk mengidentifikasi setiap *record* secara spesifik, serta relasi antar tabel yaitu penghubung antara dua atau lebih tabel melalui kunci utama dan kunci asing (*foreign key*), memungkinkan normalisasi dan pengelolaan data yang efisien.

Pada penelitian ini, *database* berfungsi sebagai sumber utama data transaksi pelanggan pada AHHAS. Data tersebut meliputi informasi pembelian, jenis barang, tanggal transaksi, dan item-item yang dibeli secara bersamaan. Melalui *database*, data transaksi disimpan dalam bentuk terstruktur sehingga dapat diekstraksi dan dianalisis menggunakan teknik *Data Mining* seperti algoritma *Apriori* dalam *RapidMiner*.

2.2 Penelitian Terdahulu

Penelitian mengenai *Data Mining*, analisis pola maupun penggunaan algoritma *Apriori* telah banyak dilakukan sebelumnya, untuk memberikan gambaran dan perbandingan mengenai penelitian ini, berikut peneliti paparkan penelitian terdahulu yang digunakan.

1. Penelitian yang dilakukan oleh Prasetya et al. (2021) berjudul "*Analisis Data Transaksi Terhadap Pola Pembelian Konsumen Menggunakan Metode Algoritma Apriori*" dan terindeks SINTA 3. Penelitian ini bertujuan untuk mengidentifikasi pola pembelian konsumen dengan memanfaatkan metode

Data Mining, khususnya algoritma *Apriori*. Penelitian ini menggunakan data penjualan dari Toko Yoss Elektronik selama periode September hingga November 2020, yang berjumlah 84 transaksi dengan 119 item. Metode yang diterapkan mencakup tahapan *Knowledge Discovery in Database (KDD)*, mulai dari *data selection*, *cleaning*, *transformation*, *Data Mining*, hingga *evaluation*. Pada tahap analisis, peneliti membandingkan dua algoritma *Association Rules*, yakni *Apriori* dan *FP-Growth*, untuk memperoleh aturan asosiasi dari data transaksi penjualan. Hasilnya, algoritma *Apriori* menghasilkan 6 aturan asosiasi dengan *confidence* tertinggi sebesar 0,64, sementara algoritma *FP-Growth* menghasilkan 25 aturan dengan *confidence* tertinggi sebesar 1,00. Penelitian ini menunjukkan bahwa data transaksi dapat dimanfaatkan untuk menemukan pola pembelian yang berguna dalam menyusun strategi promosi dan pengelolaan stok barang di toko. Perbandingan antara dua algoritma juga memberikan wawasan tentang efektivitas dan efisiensi masing-masing metode dalam menemukan pola asosiasi yang relevan dalam data transaksi penjualan (Prasetya et al., 2022).

2. Penelitian yang dilakukan oleh Merliani et al. (2022) yang berjudul “*Penerapan Algoritma Apriori Pada Transaksi Penjualan Untuk Rekomendasi Menu Makanan dan Minuman*” dan terindeks SINTA 3. Penelitian ini bertujuan untuk menerapkan algoritma *Apriori* dalam menganalisis data transaksi penjualan guna memberikan rekomendasi menu makanan dan minuman di Warung Tenda wilayah Baturaden. Landasan penelitian ini adalah kurangnya pemanfaatan data transaksi sebagai dasar

strategi promosi di tengah persaingan bisnis kuliner. Data diambil dari nota transaksi dan dianalisis menggunakan metode *Association Rule mining* dengan bantuan perangkat lunak *RapidMiner*, menggunakan nilai minimum support 20% dan confidence 70%. Hasilnya, kombinasi Es Teh Manis dan Mendoan menjadi pasangan item yang sering muncul dengan support 50% dan confidence 76%. Penelitian ini menunjukkan bahwa algoritma *Apriori* dapat membantu pelaku usaha dalam memahami pola pembelian konsumen dan menyusun strategi menu yang lebih efektif (Merliani et al., 2022).

3. Penelitian yang dilakukan oleh Vidiya dan Testiana (2023) yang berjudul “*Analisis Pola Pembelian di Lathansa Cafe & Ramen dengan Menggunakan Algoritma FP-Growth Berbantuan RapidMiner*” dan terindeks SINTA 4, bertujuan untuk menganalisis pola pembelian konsumen di sebuah kafe dengan memanfaatkan algoritma *FP-Growth*. Penelitian ini berlandaskan oleh pentingnya pemahaman terhadap perilaku pembelian pelanggan untuk menunjang strategi pemasaran dan pengelolaan stok yang lebih baik. Penelitian ini menggunakan data transaksi selama satu bulan di Lathansa Cafe & Ramen dan mengolahnya menggunakan *RapidMiner*. Hasil analisis menghasilkan aturan asosiasi dengan nilai *support* sebesar 32% dan *confidence* sebesar 94,1%, salah satunya menunjukkan bahwa pelanggan yang membeli Martabak Sayur juga pasti membeli Paket Hemat A. Temuan ini memberikan wawasan berharga bagi pemilik usaha dalam menyusun strategi menu dan promosi yang lebih efektif (Vidiya & Testiana, 2023).

4. Penelitian yang dilakukan oleh Diky Juliadi, Bambang Irawan, Agus Bahtiar, dan Odi Nurdiawan (2023) yang berjudul “*Penerapan Algoritma FP-Growth dan Association Rules pada Pola Pembelian Pizza Hut*” dan diterbitkan dalam *JATI (Jurnal Mahasiswa Teknik Informatika)* Vol. 7 No. 6 (SINTA 5), bertujuan untuk menganalisis pola pembelian konsumen pada salah satu restoran cepat saji, yaitu Pizza Hut, menggunakan pendekatan algoritma *FP-Growth* dan *Association Rules* berbantuan perangkat lunak *RapidMiner*. Penelitian ini berlandaskan oleh kebutuhan pelaku usaha dalam menggali informasi tersembunyi dari data transaksi penjualan guna menyusun strategi stok dan promosi yang lebih efektif. Dengan menggunakan data menu dan transaksi dari Pizza Hut serta parameter minimum *support* sebesar 20% dan minimum *confidence* sebesar 70%, penelitian ini berhasil menemukan dua aturan asosiasi yang signifikan. Salah satunya menunjukkan bahwa pelanggan yang membeli Pan Pizza cenderung juga membeli *Skinny Slice*, dengan nilai *support* sebesar 41% dan *confidence* sebesar 73%. Hasil ini diharapkan dapat dimanfaatkan oleh pihak manajemen dalam menyusun menu *bundling* dan promosi penjualan secara lebih terarah (Juliadi et al., 2024).
5. Penelitian yang dilakukan oleh Paisal Soleh, Abu Tholib, dan M. Noer Fadli Hidayat (2021) yang berjudul “*Penerapan Data Mining untuk Analisa Pola Pembelian Produk Menggunakan Algoritma Frequent Pattern – Growth*” dan diterbitkan dalam *Jurnal Rekayasa* Vol. 14 No. 3 (SINTA 3), bertujuan untuk menganalisis pola pembelian konsumen di toko Avindo Motor menggunakan

algoritma *FP-Growth*. Landasan penelitian ini adalah kebutuhan toko untuk mengelola ratusan data transaksi harian yang sebelumnya belum dimanfaatkan secara maksimal. Memanfaatkan pendekatan *frequent itemset* dan asosiasi aturan, peneliti berhasil menemukan pola hubungan antar produk yang sering dibeli bersamaan. Hasil penelitian menunjukkan beberapa aturan asosiasi yang relevan, seperti: jika pelanggan membeli mur maka mereka cenderung juga membeli clip, dengan *support* sebesar 25%, *confidence* 42,59%, dan *lift* 1,08; serta pola lainnya antara pembelian ban dalam, mur, dan clip. Penelitian ini memberikan manfaat strategis bagi manajemen toko dalam menyusun promosi dan penyediaan stok produk secara lebih efisien berdasarkan kebiasaan belanja pelanggan (Soleh et al., 2022).

6. Penelitian yang dilakukan oleh Arief Setyo Widodo dan Anief Fauzan Rozi (2021) yang berjudul “*Analisis Pola Pembelian di Kedai Bang Yhoga’s Menggunakan Metode Apriori*” dan diterbitkan dalam *Jurnal Sistem Informasi dan Bisnis Cerdas (SIBC)* Vol. 14 No. 1 (SINTA 5), bertujuan untuk menganalisis pola pembelian konsumen di sebuah kedai makanan menggunakan algoritma *Apriori* berbantuan aplikasi Weka. Penelitian ini dilatarbelakangi oleh meningkatnya persaingan bisnis kuliner di Yogyakarta dan perlunya strategi pemasaran yang lebih efisien. Dengan mengolah 819 data transaksi dari bulan April hingga Juni 2019, peneliti menerapkan tahapan *Data Mining* mulai dari *preprocessing* data fisik hingga transformasi ke format CSV untuk dianalisis di Weka. Hasil analisis menunjukkan terdapat 50 kombinasi item dengan nilai *confidence* minimal 90%, dan di antaranya

terdapat kombinasi dengan *confidence* sempurna (1.0), seperti pembelian *Coffee* dan *Original Tea Cheese* yang selalu disertai dengan *Orange Puch*. Penelitian ini menghasilkan 30 rekomendasi menu utama bagi pemilik kedai, yang dapat digunakan sebagai dasar dalam strategi menu dan promosi yang lebih tepat sasaran (Widodo et al., 2021).

7. Penelitian yang dilakukan oleh Agnes Eka Noviyanti dan Safitri Juanita (2024) yang berjudul “*Rekomendasi Paket Pakaian Berdasarkan Pola Penjualan Menggunakan Algoritma Apriori*” dan diterbitkan dalam Jurnal SISFOTENIKA Vol. 14 No. 2 (SINTA 5), bertujuan untuk menganalisis pola penjualan produk pakaian wanita di Toko Alys Studio dengan menggunakan algoritma *Apriori* dan perangkat lunak WEKA 3.8. Penelitian ini berlandaskan pada tingginya persaingan industri *fashion* serta kebutuhan pemilik toko untuk menemukan strategi *bundling* produk yang dapat meningkatkan minat beli pelanggan. Melalui pengolahan dari 885 data transaksi penjualan dari bulan Juni hingga September 2021, penelitian ini menggunakan metode CRISP-DM yang mencakup tahap pemahaman bisnis, persiapan data, pemodelan, hingga evaluasi dan deployment. Hasil pemodelan menunjukkan adanya 17 aturan asosiasi yang signifikan, terdiri dari 10 pola kombinasi 2 produk dan 7 pola kombinasi 3 produk, dengan nilai minimum support 15% dan confidence 50%. Salah satu pola terbaik yang ditemukan adalah kombinasi produk Aly, Aquarius, dan Aqila (BED) dengan *confidence* 88% dan *lift ratio* 1.98. Penelitian ini memberikan kontribusi nyata bagi strategi *bundling* produk pakaian di toko ritel dengan

memanfaatkan teknik *Data Mining* untuk pengambilan keputusan berbasis data (Noviyanti & Juanita, 2024).

8. Penelitian yang dilakukan oleh Intan Pratiwi, Nana Suarna, dan Tati Suprapti (2024) yang berjudul "*Implementasi Association Rules Mining untuk Analisis Pola Pembelian Paket Kuota Perdana Pelanggan XL Menggunakan Algoritma Apriori*" dan diterbitkan dalam JATI (Jurnal Mahasiswa Teknik Informatika) Vol. 8 No. 1 (SINTA 4), bertujuan untuk mengidentifikasi pola pembelian paket kuota perdana oleh pelanggan PT. XL Axiata Tegal. Permasalahan utama penelitian ini adalah belum adanya analisis mendalam terhadap kebiasaan pembelian pelanggan yang dapat digunakan untuk optimalisasi strategi pemasaran. Penelitian ini menggunakan 22.121 data transaksi pelanggan pada bulan Agustus 2023, peneliti menerapkan metode *Association Rules Mining* dengan algoritma *Apriori* melalui software RapidMiner. Setelah melakukan tahapan KDD (*Knowledge Discovery in Databases*), peneliti menetapkan ambang batas minimum *support* sebesar 30% dan *confidence* 60%. Hasilnya, ditemukan tujuh aturan asosiasi dengan tingkat *confidence* tertinggi sebesar 93%, seperti pembelian *PV Hotrod Special M (4GB)* yang sering disertai dengan *PV Hotrod Special S (2.5GB)*. Penelitian ini memberikan wawasan strategis bagi perusahaan dalam menentukan kombinasi produk unggulan serta estimasi produksi dan distribusi stok paket kuota (Pratiwi et al., 2024).
9. Penelitian yang dilakukan oleh Andreas Lewis, Muhammad Zarlis, dan Zakarias Situmorang (2021) berjudul "*Penerapan Data Mining*

Menggunakan Task Market Basket Analysis pada Transaksi Penjualan Barang di AB Mart dengan Algoritma Apriori” yang diterbitkan dalam Jurnal Media Informatika Budidarma Vol. 5 No. 2 (SINTA 3), bertujuan untuk menerapkan teknik *Market Basket Analysis* (MBA) guna menemukan pola pembelian konsumen di AB Mart. Permasalahan yang diangkat dalam penelitian ini berkaitan dengan penataan barang yang tidak memperhatikan kebiasaan konsumen dalam membeli produk secara bersamaan, sehingga berdampak pada efektivitas layanan toko. Menggunakan *software* Weka dan algoritma *Apriori*, penelitian ini memproses data transaksi dari bulan Agustus dan Oktober 2018, serta menetapkan nilai *support* minimum sebesar 0,1 dan *confidence* minimum 0,9. Hasil analisis menemukan beberapa kombinasi produk yang sering dibeli bersamaan, seperti Pepsodent, Frisian Flag, dan Indomilk, dengan *confidence* tertinggi sebesar 90%. Penelitian ini menyimpulkan bahwa penerapan *Data Mining* dapat memberikan rekomendasi strategis bagi penataan produk untuk meningkatkan kenyamanan belanja dan efektivitas penjualan (Lewis et al., 2021).

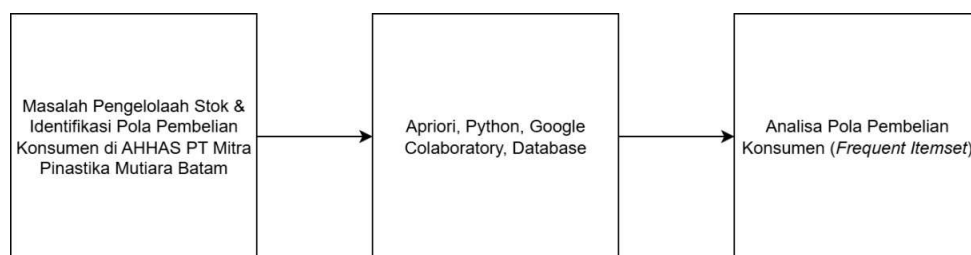
10. Penelitian yang dilakukan oleh Johannes Tarigan, Zainuddin, dan Rika Safrina (2022) yang berjudul “*Penerapan Algoritma Apriori dalam Menentukan Pola Pembelian Barang Konsumen pada PT. Indo April Medan*” dan diterbitkan dalam Jurnal Mantik Vol. 6 No. 2 (SINTA 5), bertujuan untuk menemukan pola pembelian konsumen guna mendukung strategi pemasaran dan peningkatan pelayanan di PT. Indo April Medan. Landasan penelitian ini adalah banyaknya data transaksi yang belum dimanfaatkan secara optimal

untuk menghasilkan informasi yang relevan dalam pengambilan keputusan. Penelitian ini menggunakan algoritma *Apriori* dengan bantuan perangkat lunak *RapidMiner* dalam menganalisis 50 data transaksi penjualan. Nilai minimum *support* yang digunakan adalah 40% dan *confidence* 60%. Hasil penelitian menghasilkan beberapa aturan asosiasi yang menunjukkan hubungan kuat antar produk, seperti konsumen yang membeli Sapu cenderung juga membeli Tempat Sampah, dengan nilai *confidence* mencapai 100%. Penelitian ini memberikan manfaat praktis bagi perusahaan dalam penyusunan strategi penempatan barang dan promosi yang lebih efektif.

2.3 Kerangka Penelitian

Kerangka penelitian ini berfokus pada analisis pola pembelian konsumen di AHHAS PT. Mitra Pinasthika Mustika Batam menggunakan metode algoritma *Association Rules*, khususnya algoritma *Apriori*. Input utama dalam penelitian ini adalah data transaksi pembelian konsumen yang mencakup item-item suku cadang dan layanan servis yang dibeli dalam satu nota transaksi. Penelitian ini bertujuan untuk mengidentifikasi keterkaitan antar produk yang sering muncul secara bersamaan dalam transaksi, yang dapat dimanfaatkan untuk mendukung pengambilan keputusan strategis seperti perancangan *bundling* produk, promosi yang tepat sasaran, serta pengelolaan stok yang lebih efisien. Untuk mengolah data tersebut, digunakan perangkat lunak *Google Colaboratory* sebagai alat bantu dalam proses *Data Mining*. Algoritma *Apriori* dipilih karena kemampuannya dalam mengekstraksi *frequent itemset* berdasarkan nilai minimum *support* dan *confidence*

yang ditentukan. Menggunakan pendekatan ini, penelitian diharapkan mampu memberikan wawasan baru bagi manajemen AHHAS dalam meningkatkan kualitas layanan dan efisiensi operasional secara *data-driven*. Di bawah ini disajikan kerangka pemikiran dalam penelitian ini.



Gambar 2. 4. Kerangka Penelitian
Sumber: (Data Penelitian, 2025)

Langkah awal dalam penelitian ini adalah mengidentifikasi permasalahan utama yang dihadapi oleh AHHAS PT. Mitra Pinasthika Mustika Batam, yaitu belum optimalnya pengelolaan stok suku cadang serta tidak adanya sistem yang mampu mengenali pola pembelian konsumen secara sistematis. Banyaknya data transaksi yang dimiliki AHHAS selama ini belum dimanfaatkan untuk mengetahui kombinasi produk atau layanan servis yang sering dibeli bersamaan. Permasalahan ini berdampak pada inefisiensi operasional, *seperti overstock dan out-of-stock*, serta strategi promosi yang belum tepat sasaran.

Untuk menjawab permasalahan tersebut, digunakan metode algoritma *Apriori* dalam pendekatan *Association Rules Mining* yang mampu menemukan *frequent itemsets* dari data transaksi pelanggan. Proses ini dibantu oleh perangkat lunak *Google Colaboratory* berbasis *Python*, yang berfungsi sebagai platform pemrosesan dan visualisasi data. Data yang dianalisis berasal dari *database*

transaksi pembelian konsumen dan layanan servis di AHHAS. Seluruh proses dilakukan melalui tahapan KDD (*Knowledge Discovery in Database*) mulai dari seleksi data, *preprocessing*, hingga *mining* dan evaluasi.

Tahap akhir dari penelitian ini adalah analisis terhadap pola pembelian konsumen yang berhasil diidentifikasi dari hasil *mining data*. Pola ini berupa aturan-aturan asosiasi yang menunjukkan hubungan antar produk atau layanan yang sering muncul dalam satu transaksi. Hasil analisis ini digunakan sebagai dasar untuk pengambilan keputusan strategis di AHHAS, seperti merancang bundling produk, memperkirakan permintaan stok, serta meningkatkan efisiensi pelayanan kepada pelanggan. Menggunakan pendekatan ini, AHHAS diharapkan dapat meningkatkan kualitas layanan dan efisiensi bisnis secara berkelanjutan berbasis data.