

BAB II

TINJAUAN PUSTAKA

2.1 Algoritma

Algoritma adalah serangkaian langkah logis dan terstruktur yang dirancang untuk menyelesaikan suatu masalah atau mencapai tujuan tertentu. Dalam konteks komputasi, algoritma berfungsi sebagai pedoman bagi komputer untuk melakukan tugas-tugas spesifik, seperti pengolahan data, pengambilan keputusan, atau perhitungan matematis.

Menurut Kamus Besar Bahasa Indonesia (KBBI), algoritma adalah "prosedur sistematis untuk memecahkan masalah matematis dalam langkah-langkah terbatas atau urutan pengambilan keputusan logis untuk memecahkan masalah tersebut." Algoritma juga dapat diterapkan di berbagai bidang, seperti matematika, ilmu komputer, ekonomi, dan teknik, untuk membantu menyelesaikan masalah yang memerlukan analisis yang logis dan efisien [5].

2.1.1 Sistem Algoritma

Sistem Algoritma merujuk pada kumpulan atau integrasi dari beberapa algoritma yang bekerja bersama dalam suatu sistem untuk menyelesaikan tugas atau masalah yang lebih kompleks. Berbeda dengan algoritma tunggal yang fokus pada penyelesaian satu tugas, sistem algoritma menggabungkan berbagai algoritma untuk menghasilkan solusi yang lebih menyeluruh.

Dalam sistem algoritma, setiap algoritma memiliki perannya masing-masing dan

saling mendukung untuk mencapai tujuan yang lebih besar. Sistem algoritma sering digunakan dalam aplikasi teknologi, seperti sistem rekomendasi, kecerdasan buatan (AI), dan pengolahan data besar [6].

2.1.2 Parameter Utama

Algoritma Apriori menggunakan dua parameter utama untuk menemukan hubungan antar item:

1. Support:

Mengukur seberapa sering itemset tertentu muncul dalam dataset. Rumus:

$$1. \text{Support (A)} = \frac{\text{Jumlah Transaksi Mengandung A}}{\text{Total Transaksi}}$$

2. Confidence:

Sedangkan untuk menemukan nilai confidence sebuah item dapat menggunakan rumus persamaan. Mengukur seberapa besar kepercayaan bahwa kehadiran item A dalam transaksi diikuti oleh item B.

$$1. \text{Confidence (A} \rightarrow \text{B)} = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Jumlah Transaksi Mengandung A}}$$

Algoritma Apriori bekerja secara iteratif, mulai dari itemset kecil hingga menghasilkan itemset yang lebih besar.

3. Rule Generation (Pembuatan Aturan Asosiasi)

Setelah menemukan itemset yang sering muncul, aturan asosiasi dibuat berdasarkan parameter *confidence*. Algoritma Apriori menggunakan prinsip ini untuk mengurangi jumlah itemset yang perlu diuji. Jika suatu itemset tidak memenuhi syarat minimum *support*, maka semua *superset*-nya juga tidak akan memenuhi syarat.

4. Hasil Akhir

Output dari algoritma Apriori berupa:

Frequent Itemsets: Item-item yang sering muncul bersama.

Association Rules: Hubungan antara item yang memiliki nilai *support* dan *confidence* yang tinggi.

2.2 Algoritma Apriori

2.2.1 Definisi Algoritma Apriori

Algoritma Apriori adalah metode dasar dalam data mining yang diperkenalkan oleh Agrawal & Srikant pada tahun 1994 untuk mengidentifikasi kumpulan item yang sering muncul dalam aturan asosiasi berbasis boolean. Teknik ini digunakan untuk menemukan keterkaitan antara berbagai item dalam suatu dataset, yang sering disebut sebagai affinity analysis atau market basket

analysis. Dengan algoritma ini, pola hubungan antar produk dapat diidentifikasi untuk membantu pengambilan keputusan dalam berbagai bidang, terutama dalam dunia bisnis dan pemasaran.

Association rule mining merupakan salah satu teknik dalam data mining yang berfungsi untuk mengungkap aturan asosiasi antara kombinasi item dalam suatu dataset. Misalnya, dalam analisis perilaku belanja di supermarket, teknik ini dapat mengungkap kemungkinan pelanggan membeli roti bersamaan dengan susu. Informasi tersebut dapat dimanfaatkan oleh pemilik toko untuk mengoptimalkan tata letak produk atau menerapkan strategi promosi, seperti pemberian diskon pada kombinasi barang tertentu guna meningkatkan penjualan.

Market Basket Analysis (analisis keranjang belanja) adalah salah satu aplikasi utama dari analisis asosiasi, yang menjadi terkenal karena penggunaannya untuk menganalisis isi keranjang belanja konsumen di pasar swalayan. Analisis asosiasi, khususnya teknik frequent pattern mining, telah menarik perhatian banyak peneliti karena kemampuannya untuk menghasilkan algoritma yang efisien dalam menemukan pola-pola frekuensi tinggi.

Keberhasilan dari suatu aturan asosiasi dapat diukur dengan dua parameter utama: support (nilai penunjang) dan confidence (nilai kepastian). *Support* mengukur persentase kombinasi item dalam database, sedangkan *confidence* mengukur kuatnya hubungan antar item dalam aturan asosiasi tersebut [7]

2.2.2 Definisi Association Rule

Analisis keranjang belanja (Market Basket Analysis) adalah contoh penggunaan teknik data mining yang dikenal sebagai *Association Rule*, yang digunakan untuk menentukan hubungan atau asosiasi antara satu set variabel atau item dalam sebuah dataset.[8].

Dalam konteks analisis data, Association Rules membantu dalam mengidentifikasi keterkaitan antar item yang dapat digunakan untuk pengambilan keputusan bisnis, seperti penempatan produk atau promosi penjualan. Misalnya, dengan mengetahui produk apa saja yang sering dibeli bersamaan, perusahaan dapat mengatur strategi pemasaran yang lebih efektif [9].

Algoritma Apriori adalah salah satu algoritma yang digunakan dalam Association Rules untuk menemukan frequent itemsets dalam dataset. Algoritma ini bekerja dengan mengidentifikasi item-item yang sering muncul bersama dan membangun aturan asosiasi berdasarkan item-item tersebut. Penggunaan algoritma Apriori memungkinkan analisis yang efisien terhadap data transaksi yang besar [10].

Algoritma Apriori terdiri dari dua proses utama, yaitu :

- a. Penggabungan: Proses ini menggabungkan semua item sampai tidak dapat terbentuk kombinasi lagi.
- b. Pruning (pemangkasan): Dalam proses ini, hasil kombinasi item akan dipotong sesuai dengan tingkat dukungan minimal yang telah ditetapkan. Algoritma Apriori dibagi menjadi beberapa tahap yang dikenal sebagai

narasi atau pass. Tahap-tahap tersebut adalah sebagai berikut: [11].

1. Pembentukan kandidat itemset. Kandidat k-itemset berasal dari kombinasi (k-1)-itemset yang diperoleh dari iterasi sebelumnya. Salah satu cara untuk menggunakan algoritma apriori adalah dengan memangkas kandidat k-itemset yang subsetnya berisi item k-1 yang tidak termasuk dalam pola frekuensi tinggi dengan panjang k-1.
2. Penghitungan support dari tiap kandidat k-itemset. Support dari tiap kandidat k-itemset dihitung dengan men-scan database untuk mengetahui jumlah transaksi yang memuat semua item.
3. Tetapkan pola frekuensi tinggi. Pola ini akan memuat k-item atau k-itemset dari kandidat k-itemset dengan support lebih besar dari minimum.
4. Bila tidak didapat pola frekuensi tinggi baru maka seluruh proses dihentikan. Penerapan Association Rules tidak hanya terbatas pada analisis keranjang belanja, tetapi juga dapat digunakan dalam berbagai bidang lain seperti analisis pola perilaku pengguna di situs web, deteksi penipuan, dan rekomendasi produk. Dengan demikian, teknik ini menjadi alat yang penting dalam data mining untuk mengungkap pola tersembunyi dalam data yang besar [12].

2.2.3 Parameter Aturan Asosiasi

Aturan asosiasi merupakan teknik untuk menemukan hubungan atau keterkaitan antara suatu item dengan item lainnya dalam suatu dataset. Metode ini sering dikenal sebagai analisis keranjang belanja (shopping basket analysis) dan bekerja dengan pola "jika-maka" untuk mengidentifikasi hubungan antar produk. Pola asosiasi ini diperoleh berdasarkan analisis probabilitas dari kumpulan data transaksi, memungkinkan bisnis memahami kecenderungan pembelian pelanggan.[13].

1. Support(s) nilai support sebuah item diperoleh dengan menggunakan rumus berikut:

$$1. \text{Support (A)} = \frac{\text{Jumlah Transaksi Mengandung A}}{\text{Total Transaksi}}$$

Support dari dua item diperoleh dari rumus berikut:

2. Confidence (c) adalah metrik yang mengukur seberapa sering item A muncul dalam transaksi yang juga mencakup item A. Perhitungan confidence dilakukan setelah menentukan nilai support. Nilai confidence dari aturan $A \rightarrow B$ dapat diperoleh menggunakan rumus berikut:

$$1. \text{Confidance } (A \rightarrow B) = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Jumlah Transaksi Mengandung A}}$$

2.3 Data Mining

2.3.1 Pengertian Data Mining

Data mining adalah proses penggalian sejumlah besar data untuk menemukan pola, hubungan, atau informasi tersembunyi yang berguna. Tujuan dari proses ini adalah untuk menghasilkan informasi dari data mentah yang dapat digunakan untuk pengambilan keputusan. Dalam konteks bisnis, data mining digunakan untuk memahami perilaku pelanggan, meningkatkan efisiensi operasional, dan merancang strategi pemasaran.

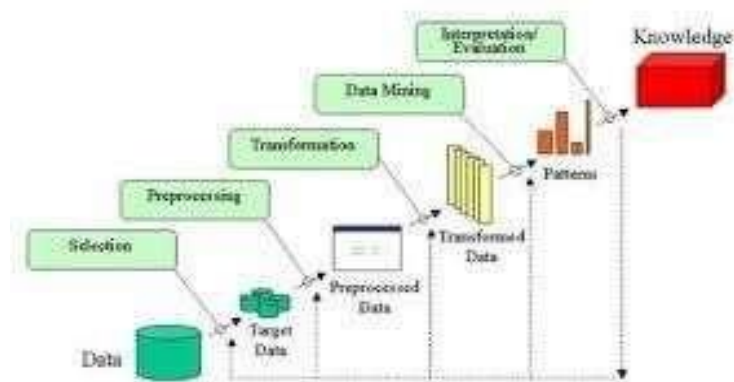
Data mining merupakan proses kumpulan data untuk mendapatkan hubungan yang tak diduga dan menyederhanakan data dengan cara yang berbeda dari sebelumnya, sehingga dapat dipahami dan memiliki manfaat bagi pemilik data [14].

Pada titik ini, proses melakukan pencarian pola, desain, dan informasi dari sebuah database melalui penggunaan berbagai metode, teknik, atau algoritma. Tujuan dan proses secara keseluruhan bergantung pada pemilihan metode yang tepat untuk proses pemerolehan informasi.

2.3.2 Knowledge Discovery In Database Knowledge Discovery In Database (KDD)

Knowledge Discovery in Database (KDD) adalah proses mencari dan mengidentifikasi pola tersembunyi dalam data yang sah, baru, bermanfaat, dan dapat dimengerti. Dengan perkembangan pesat teknologi dan informasi, data menjadi elemen penting dalam bisnis dan pengambilan keputusan, memungkinkan pengusaha memanfaatkan informasi untuk mendukung strategi dan keberlanjutan usaha.

Dalam proses Knowledge Discovery in Database (KDD) seleksi data merupakan proses himpunan data dan menciptakan data target pada sampel data, dimana penemuan akan dilakukan selanjutnya hasil seleksi [15]. Knowledge Discovery in Database (KDD) adalah salah satu metode data mining yang mencakup berbagai tahapan dalam mengubah data menjadi pengetahuan bermanfaat. Seringkali dianggap sebagai sinonim dengan data mining, KDD sebenarnya adalah konsep yang berbeda tetapi saling terkait. KDD adalah proses lengkap yang mencakup berbagai langkah, seperti pemilihan data, pra-pemrosesan, transformasi, data mining, dan akhirnya evaluasi dan interpretasi hasilnya. Namun, data mining adalah bagian penting dari KDD, di mana teknik dan algoritma tertentu digunakan untuk mengekstrak pola atau informasi dari data yang telah diproses.



Gambar 2.1 Ilustrasi Proses Knowledge Discovery In Database

Untuk menemukan dan mengenali pola dalam suatu database, proses Knowledge Discovery in Database (KDD) digunakan. Beberapa tahapan utama terdiri dari proses ini, seperti :

1. Seleksi Data

Pada tahap ini, data dari database utama dipilih untuk digunakan dalam analisis. Kumpulan data target yang akan digunakan dalam pencarian pola akan dibentuk dari data yang telah dipilih. Untuk memudahkan analisis lebih lanjut, hasil pilihan ini kemudian disimpan dalam berkas yang berbeda dari database operasional.

2. Pemrosesan dan Pembersihan Data

Tujuan dari proses ini adalah untuk menghilangkan data yang tidak relevan, duplikat, dan ketidakkonsistenan dari dataset. Sebelum data dimindahkan, mereka harus dibersihkan. Proses pembersihan ini mencakup pemeriksaan kesalahan, penghapusan data yang tidak akurat, dan penyesuaian untuk memastikan bahwa data

tersebut lebih baik. Selain itu, proses ini juga melibatkan peningkatan menambahkan informasi tambahan yang relevan untuk membuat analisis lebih menyeluruh.

3. Transformasi Data

Transformasi adalah proses pengolahan ulang data yang telah dipilih agar sesuai dengan kebutuhan analisis data mining. Proses ini bersifat fleksibel dan bergantung pada jenis pola atau informasi yang ingin ditemukan dalam database.

4. Data Mining

Pada tahap ini, metode atau teknik tertentu diterapkan untuk mengidentifikasi pola atau informasi yang tersembunyi dalam data. Berbagai algoritma dapat digunakan sesuai dengan tujuan analisis dan kebutuhan proses KDD secara keseluruhan. Pemilihan metode bergantung pada jenis data serta hasil yang ingin diperoleh.

5. Evaluasi dan Interpretasi

Tahap akhir dari KDD adalah mengevaluasi hasil yang diperoleh dari proses data mining. Hasil yang ditemukan akan diperiksa apakah sesuai dengan fakta atau hipotesis sebelumnya. Informasi yang diperoleh kemudian diterjemahkan ke dalam format yang lebih mudah dipahami agar dapat digunakan secara efektif oleh pihak yang berkepentingan.

2.3.3 Proses Data Mining

Tujuan proses data mining adalah untuk menghasilkan informasi yang relevan untuk pengambilan keputusan. Menurut Putri (2024), langkah-langkah dalam proses

data mining meliputi [16]:

1. Pemahaman Bisnis (Business Understanding)

Tahap awal ini bertujuan untuk memahami tujuan atau masalah bisnis yang akan diselesaikan menggunakan data mining. Misalnya, pada data penjualan ayam, tujuan dapat berupa menemukan pola pembelian konsumen.

2. Pemahaman Data (Data Understanding)

Mengumpulkan dan mengeksplorasi data yang tersedia untuk memahami struktur, kualitas, dan kelengkapan data. Tahap ini membantu mengidentifikasi data yang relevan untuk dianalisis.

3. Persiapan Data (Data Preparation)

Tahapan ini melibatkan pembersihan data (data cleansing), transformasi data, dan format ulang data agar siap digunakan dalam algoritma data mining. Langkah ini sangat penting untuk memastikan akurasi hasil analisis.

4. Pemodelan (Modeling)

Menerapkan algoritma data mining, seperti algoritma Apriori, untuk menganalisis data dan menghasilkan pola atau hubungan. Pada tahap ini, parameter algoritma diatur agar sesuai dengan karakteristik data.

5. Evaluasi (Evaluation)

Mengevaluasi model yang dihasilkan untuk memastikan bahwa pola atau hubungan yang ditemukan sesuai dengan tujuan bisnis dan memiliki tingkat akurasi yang tinggi.

6. Penerapan (Deployment)

Mengimplementasikan hasil analisis ke dalam sistem atau proses bisnis. Pada tahap ini, pola yang ditemukan digunakan untuk pengambilan keputusan atau strategi pemasaran. Proses ini, seperti yang disebutkan oleh Yustina dan Surya (2022), bersifat interaktif, dimana analisis dapat meningkatkan akurasi dan relevansi hasil [15].

Data mining adalah proses menganalisis dan mengekstraksi pola atau informasi berharga dari sekumpulan data yang besar. Teknik ini memungkinkan pengolahan data secara otomatis untuk menemukan hubungan tersembunyi, tren, atau pola tertentu yang dapat digunakan dalam pengambilan keputusan. Data mining sering diterapkan di berbagai bidang, seperti bisnis, kesehatan, keuangan, dan pemasaran, guna meningkatkan efisiensi serta memahami perilaku pelanggan atau tren pasar. Dalam penerapannya, data mining menggunakan berbagai metode, termasuk klasifikasi, klusterisasi, asosiasi, dan prediksi untuk menggali wawasan yang lebih mendalam. Dengan memanfaatkan algoritma tertentu, data yang sebelumnya tidak terstruktur dapat diolah menjadi informasi yang lebih bermakna. Hasil dari proses ini membantu perusahaan dan organisasi dalam mengoptimalkan strategi, mengurangi risiko, serta meningkatkan daya saing berdasarkan analisis data yang lebih akurat.

2.4 Perangkat Lunak Pengembangan Sistem

2.4.1 Altair AI Studio



Gambar 2.2 Aplikasi Altair AI Studio

Altair AI Studio adalah perangkat lunak sumber terbuka untuk analisis data termasuk penambangan data, penambangan teks, dan analisis prediktif.. Altair AI Studio juga sebuah aplikasi atau software yang berfungsi untuk mempelajari tentang data mining untuk analisis data, pemrosesan data, dan pengembangan model prediktif dan bersifat open source [17].

2.4.2 Microsoft Excel



Gambar 2.3 Aplikasi Microsoft Excel

Microsoft Excel adalah aplikasi pengolahan data yang menjadi bagian dari paket Microsoft Office. Excel digunakan secara luas untuk pengelolaan data, analisis statistik, dan visualisasi data. Aplikasi ini memungkinkan pengguna untuk membuat tabel, grafik, serta menggunakan fungsi-fungsi matematika, logika, dan statistik

untuk menganalisis data. Dalam konteks data mining sederhana, Excel dapat digunakan untuk eksplorasi awal data, pembersihan data, dan persiapan data sebelum dianalisis dengan perangkat lunak lain seperti Altair AI Studio.

Kombinasi penggunaan Altair AI Studio dan Excel memungkinkan proses analisis data yang lebih efisien, mulai dari persiapan data hingga visualisasi hasil akhir. Excel digunakan untuk praproses data, seperti membersihkan dan memformat data transaksi, sementara Altair AI Studio digunakan untuk analisis lebih lanjut, seperti menemukan pola pembelian menggunakan algoritma Apriori.

2.5 State Of Nature

No.	Judul	Nama Penulis	Tahun	Hasil
1.	Sistem Keputusan Kinerja Dosen Absensi Data Menggunakan Algoritma Apriori Studi Kasus Fakultas Sains Dan Teknologi Universitas Labuhanbatu	Budianto Bangun, Rahmadani Pane , Ali Akbar Ritonga, Iwan Purnama, Laila Hamkimi Siregar	2020	Algoritma Apriori dapat digunakan dalam data mining untuk mengevaluasi kinerja karyawan dan menentukan rekomendasi berdasarkan pola yang ditemukan.
2.	Penerapan Data Mining Klasifikasi Kepuasan Pelanggan Transportasi Online Menggunakan Algoritma C4.5	Ely Jannah , Volvo Sihombing, Masrizal	2023	hasil klasifikasi kepuasan pelanggan pengguna transportasi online, jumlah pelanggan pengguna transportasi online dengan klasifikasi “Tidak Puas” sebanyak 91 orang pelanggan, sedangkan yang “Tidak Puas” sebanyak 9 orang pelanggan.

3.	Memanfaatkan Algoritma Apriori: Aplikasi Berbasis Web untuk Penambangan Aturan Asosiasi	Rasid Siddik1,Angga Putra Juledi, Volvo Sihombing	2024	Aplikasi web yang dibuat memiliki kemampuan untuk melakukan penambangan pola asosiasi dengan cara yang efektif dan efisien, dengan memberikan hasil yang relevan dan bermanfaat untuk membantu pengguna memahami lebih lanjut tentang data transaksional.
4.	Penerapan Data Mining Untuk Evaluasi Data Penjualan Menggunakan Metode Clustering Dan Agoritma Hirarki Divisive Studi Kasus Toko Sembako Pujo	Ade Eka Febriyanti1 , Syaiful Zuhri Harahap , Masrizal	2024	Data mining membantu mengumpulkan informasi tentang pelanggan setia yang ada di Toko Sembako Pujo. Setelah dievaluasi dengan menggunakan algoritma hirarki divisive dan diproses dengan software rapid miner, ditemukan bahwa pelanggan setia berada di cluster 3 dengan rentang 117-358.
5.	Implementasi Data Mining Algoritma Apriori untuk Meningkatkan Penjualan Produk di Lakoe Dessert Pondok Kacang	Badrul Munir	2020	mengidentifikasi pola pembelian pelanggan yang dapat digunakan untuk menghasilkan strategi pemasaran yang lebih efisien.

6.	Implementasi Data Mining pada Penjualan Seprai Menggunakan Algoritma Apriori	Asmara, R., & Saputra, D	2021	Ditemukan pola pembelian pelanggan yang dapat membantu toko membuat rencana penjualan dan promosi.
7.	Penerapan Data Mining Menggunakan Algoritma Apriori dalam Menentukan Pola Penjualan pada Toko Bangunan	Dongga, J., et al	2020	Mengidentifikasi barang-barang yang sering dibeli secara bersamaan membantu toko dalam mengatur stok dan merencanakan strategi pemasaran.
8.	Penerapan Algoritma Apriori untuk Mencari Pola Transaksi Penjualan Berbasis Web pada Cafe Sakuyan Side	T Kurniana , A Lestari , E D Oktaviyani	2023	Manajemen kafe dapat menggunakan aturan asosiasi yang kuat ini saat mereka membuat strategi penjualan.
9.	Analisis Data Mining Menggunakan algoritma apriori Pada Sistem Informasi	Yuli DestrianaSirait	2023	Penelitian ini menghasilkan kombinasi aturan asosiasi yang terbentuk dan lima iterasi. Data ini dapat digunakan untuk memberi tahu pemilik toko untuk

	Aplikasi Penjualan (Studi Kasus : PT. Pasar Swalayan Maju Bersama Letda Sujono)			mengembangkan strategi penjualan dan membentuk pola penjualan baru yang lebih baik agar penjualan di toko tersebut dapat berjalan secara maksimal.
10.	Implementasi Data Mining Algoritma Apriori untuk Meningkatkan Penjualan	Abdul Harist Nst, Ibnu Rasyid Munthe, Angga Putra Juledi	2021	Produk Pudding Strawberry dan Pudding Vanilla menjadi item yang paling sering dibeli oleh pelanggan. Mengetahui produk terlaris dan pola pembelian konsumen dapat membantu bisnis dalam merancang strategi penjualan yang lebih efektif.
11.	Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan	Robby Takdirillah	2020	Algoritma Apriori dapat dimanfaatkan untuk menganalisis data transaksi penjualan guna menemukan hubungan antar produk berdasarkan pola pembelian pelanggan. Dengan bantuan tools Orange, algoritma ini mengidentifikasi keterkaitan antar produk, memungkinkan bisnis untuk memahami pola belanja dan mengoptimalkan strategi penjualan.