

BAB II

KAJIAN PUSTAKA

2.1 Kajian Pustaka

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

1. Berdasarkan penelitian yang telah dilakukan terlebih dahulu oleh Charles Parmonangan Hutabarat dan Guntoro pada 2021 dengan judul “Penerapan Data Mining *Association Rule* Menggunakan Algoritma *FP-Growth* Untuk Persediaan *Sparepart* pada Bengkel” *Jurnal Komtika (Komputasi dan Informatika)*, Vol. 5 No. 2. Hasil penelitian menyimpulkan bahwa pencarian aturan asosiasi menggunakan algoritma *FP-Growth* menghasilkan aturan asosiasi dengan nilai *support* terendah dan paling percaya diri sebagai nilai referensi. Berdasarkan temuan, ada tiga kategori suku cadang yang sering ditawarkan pada tahun 2020 dan 2021, yaitu OM (Oli Mesin), AK (Baterai), dan BS (Baterai) (*Spark plug*). [4]
2. Berdasarkan penelitian yang telah dilakukan terlebih dahulu oleh Andriani Nadia Dwi pada tahun 2019 dengan judul “Implementasi Algoritma *Fp-Growth* dalam *Market Basket Analysis* untuk Menganalisis Pola Belanja Konsumen Pada Data Transaksi” *Jurnal Ilmiah*. Hasil penelitian menyimpulkan 17 yaitu jika membeli ROMA WAFER CHOCO BLAST 54 maka membeli TARO SW 10G dengan nilai *support* = 0.03 dan nilai *confidence* = 1. [3]
3. Berdasarkan penelitian yang telah dilakukan terlebih dahulu oleh Vincent Jessfry dan Muhammad Siddik pada tahun 2024 dengan judul “Penerapan Data Mining Menggunakan Algoritma Apriori Dalam Membangun Sistem Persediaan” *JOISIE (Journal Of Information Systems And Informatics Engineering)* Vol. 8, No.1, Juni 2024 ISSN: 2527-3116 Hasil penelitian menyimpulkan bahwa berdasarkan data algoritma Apriori juga dapat melihat kombinasi barang yang saling berhubungan dimana dihasilkan 4

rule dimana *itemset* pell dan sapu dengan *confidence* 57%, *itemset* sapu dan pell dengan *confidence* 72,4%, *itemset* pell dan sikat lantai dengan *confidence* 55,2%, dan *itemset* sikat lantai dan pell dengan *confidence* 85,8%. [5]

4. Berdasarkan penelitian yang telah dilakukan terlebih dahulu oleh Rahmad Aditiya dan Sarjon Defit pada tahun 2020 dengan judul “Prediksi Tingkat Ketersediaan *Stock* Sembako Menggunakan Algoritma *FP-Growth* dalam Meningkatkan Penjualan” Jurnal Informatika Ekonomi Bisnis Vol. 2 No. 3, September 2020 ISSN: 2714-8491 Hasil penelitian menyimpulkan bahwa diperoleh 17 pola penjualan sembako yang dihitung secara manual dan ada 16 pola penjualan sembako yang dihitung menggunakan aplikasi Rapidminer 9.4. Dari pola-pola tersebut dapat direkomendasikan kepada pemilik Toko UD. [6]
5. Berdasarkan penelitian yang telah dilakukan terlebih dahulu oleh Muhammad Yusuf ML, Fairuz Azmi dan Ratna Astuti Nugrahaeni pada tahun 2023 dengan judul “Sistem Rekomendasi Penyediaan Stok Barang Berdasarkan Anggaran Pada Studi Kasus Toko Ud Rahmat Yh Banda Aceh” *e-Proceeding of Engineering* : Vol.10, No.1 Februari 2023 ISSN : 2355-9365. Hasil penelitian menyimpulkan bahwa Mayoritas pengguna beranggapan aplikasi ini dapat membantu melayani pelanggan (96%), dan aplikasi ini mudah untuk dipahami (96%), tampilan keuangan aplikasi ini dianggap memenuhi kebutuhan pengguna (92%), aplikasi juga dianggap membantu mencatat transaksi (96%), mempermudah pengelolaan stok (90%), membantu memesan stok barang (84%), dan membutuhkan waktu pelatihan yang sebentar (90%) namun tampilan aplikasi ini dinilai kurang menarik oleh sebagian pengguna (58%). [7]

Penelitian ini sejalan dengan studi sebelumnya yang menerapkan algoritma *FP-Growth* untuk analisis pola pembelian. Namun, penelitian ini menonjol dalam penerapannya pada sistem rekomendasi stok barang, sedangkan kebanyakan penelitian sebelumnya lebih terfokus pada strategi pemasaran. Penggunaan algoritma *FP-Growth* untuk pengelolaan stok memberikan kontribusi baru dalam literatur terkait manajemen persediaan.

2.2 Landasan Teori

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

2.2.1 Pengertian Data Transaksi

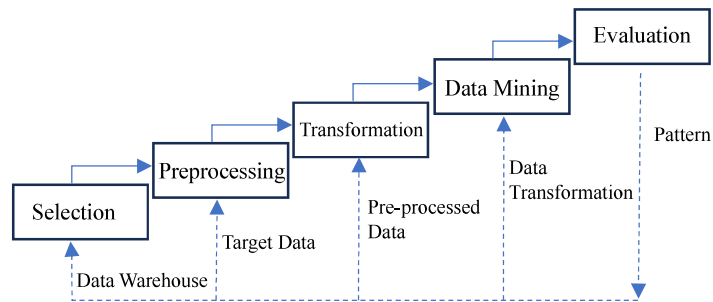
Data transaksi merujuk pada data yang mencatat peristiwa-peristiwa yang telah terjadi dalam suatu usaha atau perusahaan, yang mencakup data penjualan, pengiriman, hutang, dan peristiwa lainnya. Data transaksi biasanya dapat dikategorikan dalam tiga kelompok berdasarkan kata kerja yang digunakan, yaitu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

Definisi transaksi secara umum mencakup setiap aktivitas yang mempengaruhi aset atau keuangan organisasi atau individu. Beberapa contoh aktivitas transaksional meliputi penjualan, pembelian, pembayaran gaji, dan sebagainya. Dalam data transaksi, terdapat manajemen transaksi yang bertugas mencatat perubahan keuangan dengan cermat menggunakan metode tertentu. Data transaksi penjualan memiliki nilai penting dalam pengambilan keputusan bisnis. [8]

2.2.2 KDD (*Knowledge Discovery Database*)

KDD (*Knowledge Discovery in Databases*) berfokus pada tahapan eksplorasi data untuk menemukan pola atau pengetahuan baru dari dataset yang besar dan kompleks. Hal ini sejalan dengan tujuan penelitian, yaitu menemukan pola pembelian konsumen melalui analisis data transaksi menggunakan algoritma *FP-Growth*. KDD lebih menitikberatkan pada proses analisis data hingga menghasilkan pola atau informasi yang bermanfaat (misalnya *frequent itemsets* dalam *FP-Growth*). Penemuan pola seperti *association rules* untuk rekomendasi stok barang sangat relevan dengan prinsip KDD.

Pada kontek ini data mining merupakan satu langkah dari proses KDD, terdapat beberapa proses seperti terlihat pada gambar di bawah ini :



Gambar 2.1. Tahapan KDD

1. Seleksi Data (*Data Selection*)

Selection (seleksi/pemilihan) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *Knowledge Discovery Database* (KDD) dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional.

2. Pemilihan Data (*Preprocessing/Cleaning*)

Proses *Preprocessing* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (*tipografi*). Juga dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

3. Transformasi (*Transformation*)

Pada fase ini yang dilakukan adalah mentransformasi bentuk data yang belum memiliki entitas yang jelas ke dalam bentuk data yang valid atau siap untuk dilakukan proses Data Mining.

4. Data Mining

Pada fase ini yang dilakukan adalah menerapkan algoritma atau metode pencarian pengetahuan.

5. Interpretasi/Evaluasi (*Interpratation/Evaluation*)

Pada fase terakhir ini yang dilakukan adalah proses pembentukan keluaran yang mudah dimengerti yang bersumber pada proses Data Mining pola informasi.

2.2.3 Data Mining

Data mining ialah sebuah proses pemahaman dan identifikasi informasi berharga dari basis data yang sangat besar memanfaatkan metode seperti statistik, matematika, kecerdasan buatan, dan pembelajaran mesin. Dalam “*Knowledge Discovery in Databases*” (KDD), terdapat berbagai pendekatan yang berbeda untuk mencari informasi atau pengetahuan. Pendekatan kuantitatif meliputi pencarian probabilistik seperti logika induktif, pencarian pola, dan analisis pohon keputusan. Selain itu, terdapat pendekatan analisis kecenderungan, deviasi, algoritma genetik, jaringan saraf tiruan, serta pendekatan kombinasi dari dua atau lebih metode sebelumnya. [9]

2.2.4 Proses Data Mining

Menurut referensi [10] proses data mining terbagi dalam tiga yaitu:

1. Eksplorasi data melibatkan berbagai kegiatan seperti membersihkan data, mengubah data, mengurangi dimensi data, memilih fitur, dan lain sebagainya.
2. Pemilihan model yang tepat dan valid merupakan tahap penting dalam pembuatan dan pengujian model. Proses tersebut melibatkan seleksi model yang paling sesuai dengan masalah yang sedang dihadapi, yang dilakukan secara kompetitif.
3. Tahap penting dalam penerapan model adalah penggunaan data baru untuk menghasilkan estimasi kasus yang relevan, sehingga dapat diuji apakah model yang dibangun dapat memberikan solusi yang tepat untuk masalah yang dihadapi.

2.2.5 *Assosiation Rules*

Association Rules adalah teknik dalam data mining yang digunakan untuk menemukan hubungan atau keterkaitan antara item dalam kumpulan data besar. Setelah *itemset* frekuen ditemukan menggunakan

algoritma seperti *FP-Growth*, langkah berikutnya adalah membangun aturan asosiasi yang mengekspresikan hubungan antara item. [11]

Aturan asosiasi memiliki format:

"Jika A terjadi, maka B juga cenderung terjadi", di mana:

1. A (*antecedent*) adalah item yang ditemukan lebih dulu.
2. B (*consequent*) adalah item yang mungkin terjadi setelah A.

Langkah-langkah utama dalam membentuk aturan asosiasi:

- a. Mengidentifikasi *Frequent Itemsets*: Algoritma seperti *FP-Growth* menemukan itemset yang sering muncul dalam dataset.
- b. Membentuk Aturan Asosiasi: Berdasarkan *itemset* frekuensi, aturan dibentuk dengan memisahkan item sebagai *antecedent* dan *consequent*.
- c. Mengukur Kekuatan Aturan: Tiga metrik utama yang digunakan untuk menilai aturan adalah:
 1. *Support*: Frekuensi kemunculan itemset A dan B bersama-sama dalam dataset.
 2. *Confidence*: Seberapa sering B terjadi jika A telah terjadi.
 3. *Lift*: Mengukur sejauh mana A dan B lebih sering terjadi bersamaan dibandingkan dengan kemunculan acak.

2.2.6 Algoritma *FP-GROWTH*

Dalam analisis pola pembelian menggunakan algoritma *FP-Growth*, dua parameter utama yang menentukan kualitas hasil analisis adalah *support* dan *confidence*. Nilai-nilai ini memiliki peran penting dalam mengidentifikasi pola itemset yang sering muncul (*frequent itemsets*) dan membentuk aturan asosiasi yang relevan.

Support digunakan untuk mengukur seberapa sering suatu itemset muncul dalam dataset transaksi secara keseluruhan. Sementara itu, *confidence* digunakan untuk mengevaluasi seberapa kuat hubungan antara item dalam suatu aturan asosiasi.

A. Panduan Menentukan Support Optimal:

a. Support rendah (1%-5%):

Cocok untuk menemukan pola item jarang tetapi signifikan, misalnya produk premium atau barang musiman.

Contoh: Analisis cross-selling barang elektronik mahal yang jarang dibeli bersama.

b. Support sedang (5%-20%):

Ideal untuk kebanyakan pola transaksi di perusahaan ritel atau grosir.

Support di rentang ini umumnya mencerminkan kombinasi produk populer yang sering dibeli bersama.

c. Support tinggi (>20%):

Digunakan untuk itemset yang sangat sering muncul, misalnya produk kebutuhan sehari-hari.

Contoh: Kombinasi produk seperti beras dan minyak goreng di supermarket.

B. Panduan Menentukan Confidence Optimal:

a. Confidence rendah (<50%):

Dianggap lemah dalam dunia bisnis, karena item Y tidak selalu muncul bersama item X.

Contoh: Kombinasi produk yang kurang signifikan atau hanya terjadi secara kebetulan.

b. Confidence sedang (50%-80%):

Menunjukkan pola asosiasi yang cukup kuat dan bisa menjadi rekomendasi awal untuk perusahaan.

Contoh: Konsumen yang membeli sepatu olahraga juga cenderung membeli kaos olahraga dengan confidence 70%.

c. Confidence tinggi (>80%):

Mengindikasikan hubungan yang sangat kuat antara dua produk atau itemset.

Pola dengan confidence tinggi lebih mudah direkomendasikan untuk strategi promosi atau pengadaan stok.

2.2.7 *Unified Modelling Language (UML)*

Unified Modelling Language (UML) “adalah keluarga notasi grafis yang didukung oleh meta-model tunggal, yang membantu pendeskripsian dan desain sistem perangkat lunak, khususnya sistem yang dibangun menggunakan pemrograman berorientasi objek (OO)”. [12]

Unified Modeling Language (UML) bahasa spesifikasi standar untuk mendokumentasikan, menspesifikasikan, menggambarkan, dan membangun sistem perangkat lunak seperti halnya pada *business modelling* dan sistem lainnya. UML tidak berdasarkan pada bahasa pemrograman tertentu. Standar spesifikasi UML dijadikan standar *defacto* oleh OMG (*Object Management Group*) pada tahun 1995. UML yang berorientasikan object mempunyai beberapa notasi standar. *Unified Modeling Language (UML)* sendiri terdiri atas pengelompokan diagram – diagram sistem menurut aspek atas sudut pandang tertentu. Diagram adalah yang menggambarkan permasalahan maupun solusi dari permasalahan suatu model. UML mempunyai 9 diagram, tapi penulis hanya menggunakan 3 diagram UML, yaitu :

1. *Use Case Diagram*
2. *Activity Diagram*
3. *Class Diagram*

2.2.8 *Visual Studio 2019*

Visual Studio 2019 adalah lingkungan pengembangan terintegrasi (IDE) yang dikembangkan oleh Microsoft untuk memfasilitasi pengembangan aplikasi lintas platform, termasuk desktop, web, dan perangkat seluler. Fitur utamanya mencakup dukungan untuk berbagai bahasa pemrograman seperti *C#*, *C++*, *Visual Basic*, dan *Python*. Salah satu keunggulan *Visual Studio 2019* adalah integrasinya dengan alat-alat Git dan Azure, serta dukungan untuk pengembangan berbasis kontainer, yang mempermudah pengelolaan proyek DevOps dan lingkungan berbasis cloud.

Selain itu, *Visual Studio* 2019 memperkenalkan *IntelliCode*, fitur kecerdasan buatan yang memberikan saran kode secara cerdas, serta alat refaktorisasi yang lebih canggih untuk mempermudah pengeditan kode. Fitur debugging dan profiling yang ditingkatkan juga memudahkan pengembang dalam mendeteksi dan memperbaiki kesalahan kode. [13]

2.2.9 *Microsoft Access*

Microsoft Access adalah sistem manajemen basis data relasional (RDBMS) yang dikembangkan oleh Microsoft, bagian dari *suite Microsoft Office*. Access digunakan untuk membuat dan mengelola database dengan cepat, menggabungkan antarmuka pengguna yang mudah digunakan dengan mesin database yang kuat. Access sering digunakan untuk aplikasi database tingkat kecil hingga menengah, di mana pengguna bisa dengan mudah mengelola dan memanipulasi data melalui antarmuka grafis yang intuitif. [14]

Beberapa fitur utama dari Microsoft Access meliputi:

1. *Tabel dan Query*: Access memungkinkan pengguna membuat tabel untuk menyimpan data serta menggunakan *query SQL* untuk memfilter dan menganalisis data.
2. *Formulir dan Laporan*: Formulir digunakan untuk memasukkan dan menampilkan data, sementara laporan digunakan untuk menghasilkan dokumen yang dapat dicetak atau diekspor.
3. *VBA (Visual Basic for Applications)*: *Microsoft Access* mendukung VBA untuk memungkinkan pengguna menulis skrip yang lebih kompleks untuk mengotomatisasi tugas atau menciptakan fungsi yang lebih canggih.
4. *Relational Database*: Access mendukung hubungan antara tabel, memungkinkan integrasi data yang lebih baik dan penggunaan fungsi relasional seperti join.

Microsoft Access digunakan oleh bisnis kecil hingga menengah serta organisasi yang memerlukan solusi manajemen data yang mudah dan terjangkau tanpa memerlukan keahlian khusus dalam database *SQL Server* atau sistem manajemen database besar lainnya.

2.2.10 Microsoft Excel

Microsoft Excel adalah perangkat lunak *spreadsheet* yang dikembangkan oleh *Microsoft* dan merupakan bagian dari suite *Microsoft Office*. Excel digunakan untuk mengelola, menganalisis, dan memvisualisasikan data melalui tabel, grafik, dan alat perhitungan canggih. Program ini memungkinkan pengguna melakukan perhitungan matematis dan statistik, menyusun data dalam berbagai format, serta membuat visualisasi data dalam bentuk grafik yang interaktif. [15]

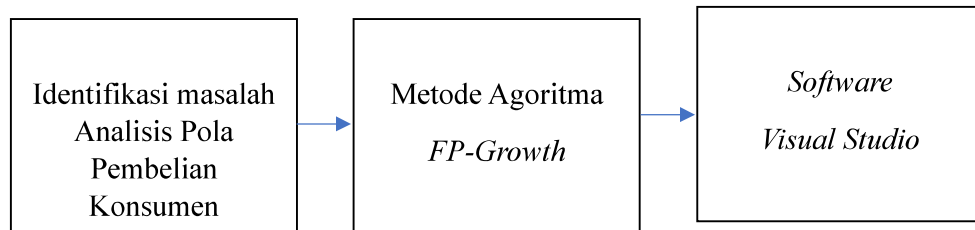
Beberapa fitur utama *Microsoft Excel* meliputi:

1. *Spreadsheet*: Excel terdiri dari sel-sel dalam grid yang diatur dalam baris dan kolom, memungkinkan pengguna memasukkan, mengorganisir, dan mengelola data dengan mudah.
2. Fungsi dan Rumus: Excel menyediakan ratusan fungsi bawaan untuk perhitungan matematika, statistik, logika, dan teks, seperti *SUM*, *AVERAGE*, *IF*, dan *VLOOKUP*. Pengguna juga dapat membuat rumus kustom.
3. *PivotTable* dan *PivotChart*: Fitur ini memungkinkan pengguna untuk menganalisis data dalam jumlah besar dengan cepat dan membuat laporan dinamis.
4. Alat Visualisasi: Excel mendukung pembuatan berbagai jenis grafik, seperti diagram batang, *pie*, *scatter*, dan garis, untuk membantu dalam penyajian data secara visual.
5. *Macros* dan *VBA (Visual Basic for Applications)*: Excel memungkinkan otomatisasi tugas melalui penggunaan *Macros* dan skrip *VBA*, membantu mempercepat proses yang berulang.
6. Kolaborasi: Versi terbaru Excel terintegrasi dengan *cloud* melalui *OneDrive* dan *SharePoint*, memungkinkan kolaborasi secara real-time dengan pengguna lain.

Microsoft Excel sering digunakan dalam berbagai bidang, termasuk akuntansi, analisis data, keuangan, serta dalam kegiatan pendidikan untuk perhitungan dan pengelolaan data yang terstruktur.

2.3 Kerangka Pemikiran

Kerangka pemikiran adalah konsepsi dari inti permasalahan yang akan dipelajari, serta menggambarkan langkah-langkah dalam proses awal hingga akhir penelitian dalam bentuk desain alur urutan penyelesaiannya. Gambar 2.2 di bawah ini merupakan kerangka pemikiran yang telah disusun oleh peneliti.



Gambar 2.2 Kerangka Pemikiran

Data yang diinput berupa masalah pada analisis pola pembelian konsumen. Data penelitian yang telah dikumpul akan dilakukan perhitungan penyelesaian menggunakan algoritma *FP-Growth*. Hasil yang diharapkan akan dilakukan pengujian dan dibuatkan aplikasi merekomendasikan data stok barang menggunakan *software visual studio 2019* dengan menggunakan bahasa *visual basic* sehinga menghasilkan produk mana saja yang memiliki nilai jual tinggi berdasarkan pola pembelian konsumen