

BAB II

TINJAUAN PUSTAKA

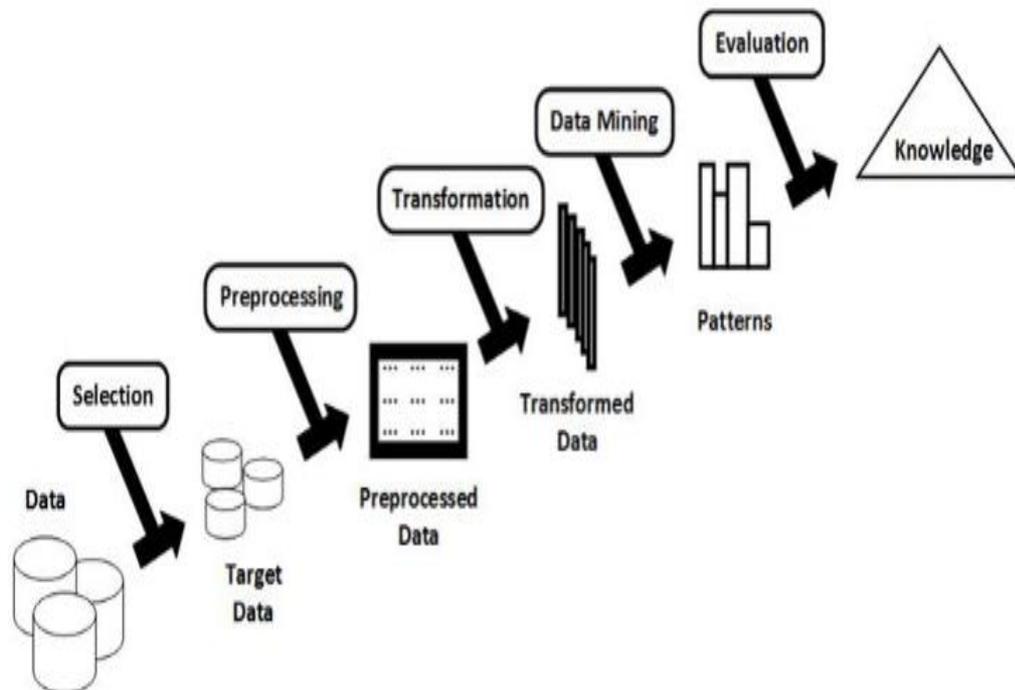
2.1 Teori Dasar

Dalam bab ini membahas tentang semua materi atau referensi dari penelitian yang dibuat oleh penulis, jurnal referensi atau buku yang dibahas yaitu *Knowledge Discovery in Database*, data mining, algoritma apriori, *software* pendukung dari data mining serta data-data yang diolah yaitu data penjualan durian di Toko Durian KimHui Batam.

2.2 Knowledge Discovery in Database (KDD)

Sebelum melakukan proses data mining akan menentukan suatu informasi didalam yang namanya *database* merupakan bagian dari *knowledge discoveri in databased* gunanya yaitu mendapatkan suatu informasi dan pola yang penting dalam data yang ada. Data yang mencari suatu informasi baru dari data yang banyak dan berguna untuk kesimpulan sebuah data yang melibatkan suatu komputer dimana manusia bisa menggunakan secara otomatis dan manual (Putri Mai Sarah Tarigan, 2022).

Knowledge Discovery in Database (KDD) salahsatu teknik yang membentuk suatu pola atau langkah dalam informasi, informasi didaptkan dalam data-data yang sangat besar yang diolah oleh penambangan data yang belum diketahui informasi dalam data tersebut (Nasir, 2020). Adapun aturan langkah dari KDD yaitu :



Gambar 2.1 langkah KDD

Penjelasan dari langka KDD yaitu menentukan data apa yang bisa diambil dari jumlah data yang besar selanjutnya diseleksi data apa yang bisa dikelola selanjutnya prosesing data yang disaring selajutnya transformasi data selanjutnya data mining menggali data menghitung dengan algoritma data mining dan selanjutnya setelang melakukan perhitungan mendapatkan hasil dari data tersebut seperti mendapatkan pengetahuan baru.

Tahapan dari proses sebelum melanjutkan proses data mining yaitu KDD, proses pahapan-tahapannya yaitu:

1. Menyeleksi data merupakan data-data yang tidak digunakan dalam proses data mining.
2. Integrasi data merupakan mengumpulkan data-data dari referensi yang berbeda.
3. Transformasi data merupakan langkah untuk mengubah data sesuai data yang bisa diproses nantinya.
4. Melakukan teknik data mining menghitung atau mencari pola yang ada.

5. Melihat pola yang dievaluasi yang ditafsirkan menjadi ilmu baru sebagai pendukung proses penarik kesimpulan.
6. Menyajikan ilmu merupakan proses penyajian yang gampang dilakukan oleh *user*.

2.3 Data Mining

Data mining bisa dikatakan suatu proses seleksi-seleksi data, eksplorasi data, dan pemodelan yang diolah dari data besar supaya mendapatkan pola dari frekuensi untuk menentukan pembelian barang (Sukmawati Anggraeni Putri, 2022). Dari defenisi lain data mining diperuntukkan untuk perhitungan matematika, kecerdasan buatan dan *machine learning* untuk mengolah suatu data yang sangat besar sehingga mendapatkan informasi yang berguna untuk pengetahuan. Data mining merupakan menambang suatu data-tata yang belum memiliki arti guna agar dari data yang bisa menghasilkan informasi yang berguna bagi penelitiannya (Koko Handoko, 2023).

Data mining adalah memudahkan untuk dimengerti untuk *database* yang dimana dalam jumlah besar dalam proses dalam menentukan suatu pola model baru yang lengkap menjabarkan bahwa mengambil suatu keputusan untuk melihat kedepannya apakah dengan data mining berisi mencari informasi yang dapat dari data yang ada dan mudah dipelajari (Piter Prasetyo Sudarto, 2023).

Data mining rangkaian langkah-langkah dari proses dalam penggalian data sehingga mendapatkan data-data penting berupa informasi, informasi yang didapat itu sangat penting untuk mendapatkan pola yang ada dari *database* (Zulham, 2021). Data mining mendefenisikan pencarian gudang data besar dari sini bisa ditemukan pola yang berguna dan menarik agar tujuan bisa memperoleh ilmu pengetahuan yang berguna dan bermanfaat (Larisma Nursinta Nainggolan, 2022).

IJBIDM provides a forum for state-of-the-art developments and research as well as current innovative activities in business intelligence, data analysis and mining. Intelligent data analysis provides powerful and effective tools for problem solving in a variety of business modelling tasks. IJBIDM highlights intelligent techniques used for business modelling, including all areas of data visualisation, data pre-processing (fusion, editing, transformation, filtering, sampling), data engineering, data mining techniques, tools and applications, neurocomputing, evolutionary computing, fuzzy techniques, expert systems, knowledge filtering, and post-processing. (Dorgham, 2022)

Menurut (Saefudin, 2019) langkah tahapan data mining meliputi :

1. Data *cleaning* (untuk menghapus data yang tidak digunakan).
2. Data *integration* (menggabungkan data dari beberapa data).
3. Data selection (data yang jelas dikembalikan atau digabungkan didalam basis data).
4. Data *transformation* (data yang sudah beda menjadi bentuk penambangan dari operasi agresi).
5. Data mining (menggunakan langkan algoritma yang ada didalam data mining).
6. Evolusi pola (mendapatkan pola dari data pengetahuan yang baru).
7. Presentasi pengetahuan (mendapatkan ilmu baru dari data yang sudah ditambah atau digali).

Menurut (Putri Mai Sarah Tarigan, 2022) data mining memiliki kelebihan dan kekurangan, kelebihan data mining yaitu:

1. Dapat menyelesaikan data besar dan data kompleks.
2. Mampu menyelesaikan data besar dari berbagai atribut dalam data.
3. Dapat mengolah data besar secara mudah dan teratur.

Kekurangan data mining yaitu :

1. Data perusahaan bisa diketahui oleh orang banyak seperti data kariawan, gaji kariawan dan menyimpan informasi penting tentang data kariawan.

2.4 Algoritma Apriori

Algoritma apriori salah satu pengambilan data dengan menggunakan langkah asosiasi yang digunakan untuk menetapkan aturan asosiasi yang ada dalam kombinasi itemset barang. Menetapkan aturan asosiasi dilakukan cara perhitungan dengan *support* dan *confidence* dari kumpulan yang ada dalam itemset. Kesimpulan yang didapat dengan nilai support tersebut lebih tinggi dari batas nilai terendahnya dan nilai *confidence* lebih tinggi dari nilai terendahnya *confidence*. Algoritma apriori sangat pas jika terdapat beberapa langkah itemset barang yang dianalisa dan sangat cocok diterapkan dalam *best seller* penjualan (Sukmawati Anggraeni Putri, 2022).

Algoritma data mining yaitu apriori mencaari suatu itemset yang dilakukan dengan assosiasi *rule* dan menggunakan ilmu dari atribut-atribut yang sudah didapatkan dalam salah satu algoritma dimana pencarian frequensi dari itemset yang ada menggunakan teknik *assosiasi rule* (Saefudin, 2019).

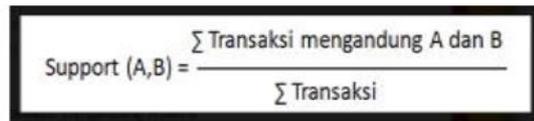
Langkah dari tahapan mencari kombinasi item dari langkah pertama untuk minimum nilai support dalam *database*, rumus dari nilai support yaitu:

$$\text{Support (A)} = \frac{\text{Jumlah transaksi mengandung A}}{\text{Transaksi Total}}$$

Gambar 2.2 Rumus Kombinasi *Itemset*

Dalam mencari support dari dua item dapat diperoleh :

$$\text{Support (A,B)} = P(A \cap B)$$



$$\text{Support (A,B)} = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi}}$$

Gambar 2.3 Rumus Nilai *Support*

Algoritma apriori merupakan proses dari pengambilan data menggunakan aturan asosiatif dari penentuan hubungan dari kombinasi item (Debora Kesia Batubara, 2020).

Menurut (Ainul Mardiah, 2020), adapun langkah-langkah algoritma apriori meliputi :

1. Analisis pola max frekuensi dan Pembentukan langkan asosiatif :

$$\text{Support (A)} = \frac{\text{Jumlah transaksi mengandung A}}{\text{Total Transaksi}}$$

$$\text{Confidence} = P(B|A) = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi mengandung A}}$$

2.5 Objek Penelitian

Durian merupakan tanaman yang berbuah, sebutan durian pertama kali dari bahasa melayu yang katanya duri diberi akhiran an supaya menjadi durian. Dari kata tersebut bisa defenisikan untuk tanaman yang kulitnya berduri-duri dan tajam. Tanaman ini didapatkan dari hutan malaysia, kalimantan, sumatra didasari oleh tanaman liar di hutan. Di indonesia di jawa disebut dengan duren, dimanado disebut duriang, di toraja disebut dulian dan seram timur disebut rulen.

Buah durian atau *durio zibethinus muur* adalah tanaman yang berasal dari tropis eksotik yang memiliki rasa yang enak dan memiliki aroma yang khas. Buah durian

disebut juga *the king of fruit* yang banyak disukai oleh kalangan masyarakat Indonesia yang memiliki rasa yang khas (Nazriah Pratiwi, 2018).

Durian yang akan dibahas di penelitian ini ada 12 jenis durian dan harga penjualan di Toko Durian KimHui:

Tabel 2.1 Jenis Durian dan Harga Durian

No	Jenis Durian	Harga per kg
1	Kampung	45.000 – 60.000 per kg
2	Musang King	230.000 – 300.000 per kg
3	Duri Hitam	350.000 – 400.000 per kg
4	Xo	180.000 – 230.000 per kg
5	Udang Merah	170.000 – 200.000 per kg
6	101	170.000 – 210.000 per kg
7	Tek Kah	150.000 – 200.000 per kg
8	D13	150.000 – 200.000 per kg
9	99	150.000 – 200.000 per kg
10	Super	110.000 – 150.000 per kg
11	Montong	180.000 – 230.000 per kg
12	D24	140.000 – 180.000 per kg

Sejarah objek penelitian ini sejak tahun 2014 Toko Durian KimHui awal mulanya beroperasi di Kota Batam tepatnya di kedai kopi budi siang malam yang berada di daerah Nagoya Batam Park, awalnya pemilik Toko Durian KimHui fokus menjual buah durian yang berasal dari Tanjung Batu dan juga dari Tanjung Pinang, dikarenakan saat itu belum ada yang menjual durian dan hanya beberapa orang yang membuka usaha durian di Kota Batam tersebut yaitu teman dari pemilik KimHui itu sendiri dan sekarang mereka masih

buka sampai saat ini mereka menjalankan bisnisnya masing-masing, dulu mereka menjual jenis durian kampung seiring berjalannya waktu Toko Durian KimHui memiliki inisiatif untuk memasukan jenis-jenis buah durian yang telah dijabarkan di atas. Mulai tahun 2020 Toko Durian KimHui Batam berpindah lokasi persis di seberang kedai kopi sebelumnya yaitu tempat makan *Seafood Acia Ikan Bakar* nagoya hingga sekarang.

2.6 *Software RapidMiner*

Dari perhitungan algoritma apriori data mining bisa dikatakan belum lengkap kalau aplikasi data mining belum digunakan, aplikasi data mining itu adalah RapidMiner. RapidMiner merupakan sebuah perangkat lunak bersifat terbuka, aplikasi ini digunakan tujuannya adalah memecahkan suatu masalah yang ada dalam data mining. Aplikasi ini memiliki lebih kurang 500 tombol prosedur operator dalam data mining, sehingga aplikasi ini banyak digunakan untuk peneliti yang menggunakan data besar seperti data mining (Yulia Andini, 2022). RapidMiner adalah hasil dari solusi agar dapat melakukan teknik deskriptif dan prediksi untuk memberi pengembangan pada pemakai agar supaya mendapatkan suatu keputusan yang benar (Fadila Shely Amalia, 2021).

Halaman utama dari *software* RapidMiner Studio versi 10.2 yaitu:



Gambar 2.4 Tampilan Utama RapidMiner Versi 10.2

2.7 Penelitian Terdahulu

Dalam penelitian ini penulis mengutip jurnal dimana jurnal tersebut membantu penulis untuk menyelesaikannya :

1. (Sukmawati Anggraeni Putri, 2022), *Analisa Pola Dalam Penentuan Best Seller Produk di FROZENIN.BUN Dengan Algoritma Apriori*. Penelitian ini menjelaskan langkah-langkah dalam melakukan *bast seller* pada barang atau produk frozen yang dijual oleh Frozenin.Bun. Dari hasil perhitungan yang dijabarkan jenis rules barang yang terjual shrimp roll maka pembeli juga membeli barang tempura dan *egg roll* nilai yang dihasilkan maksimal atau tertinggi *support* 41% dan *confidence* 63%. Jadi barang ini adalah barang yang sering terjual, sehingga pelaku usaha bisa melihat bagaimana kedepannya memiliki peluang untuk mengembangkan usahanya dan menjual barang dengan keuntungan yang tinggi.
2. (Reza Alfianzah, 2020), *Implementation of Apriori Algorithm Data Mining for Increase Sales*. Penelitian ini menjelaskan dimana perusahaan yang bertahan bisa bertahan lama dan membuat strategi yang tepat. Data barang penjualan lakoe dessert pondok kacang dari sini akan menghasilkan data yang tertumpuk dan tidak digunakan maka sayang rasanya jika tidak di mining data tersebut. Algoritma apriori dan penerapan menggunakan RapidMiner dengan nilai dukungan 2,4% dan nilai kepercayaan 50%, dari sini terlihat pembeli barang yang sering terjual.
3. (Zulham, 2021), *Analisa Pola Sistem Penjualan Makanan Ringan dengan Menggunakan Algoritma Apriori*. Penelitian ini menjelaskan penjualan tentang makanan ringan dimana banyak dibutuhkan oleh masyarakat, penjual menjual barangnya ke masyarakat sesuai permintaan kebutuhan dari pembeli dan harganya

sesuai dengan barangnya tersebut. Dari barang makanan ringan tersebut produsen mengalami kesulitan dalam permintaan pembeli sehingga konsumen mencari solusi agar makanan ringan ini selalu ada dan tidak kehabisan stoknya agar pembeli tidak mencari konsumen lain.

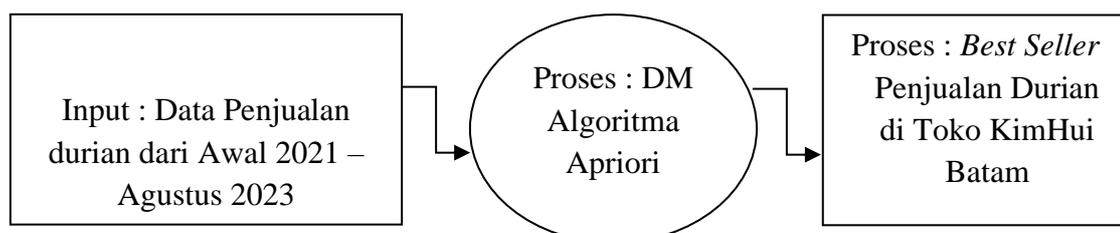
4. (Saefudin, 2019), Penerapan Data Mining Dengan Metode Algoritma Apriori Untuk Menentukan Pola Pembelian Ikan. Penelitian ini menjelaskan kegiatan pembelian ikan disini masih menggunakan teknik manual seperti mencatat dibuku catatan semua transaksi penjualan ikan, pendapatan ikan dan masih memprediksi berapa ikan yang akan dibeli untuk persediaan digudang. Dalam mendukung kegiatan atau mencari pola dalam penelitian data mining penulis menghitung dengan algoritma apriori bagaimana hasil yang akan dilihat jenis ikan mana yang yang paling diminati oleh konsumen.
5. (Delpia Guci, 2021), Implementasi Data Mining Dalam Memperoleh Pola Kombinasi Produk *Ice Cream* Di Luigi Gelato Menggunakan Algoritma Apriori. Penelitian ini menjelaskan masalah dari penjualan *ice cream* di lugi galato supaya membuat suatu strategi-stratei penjualan atau promosi *ice crime* tersebut menggunakan ilmu baru seperti kecerdasan buatan yaitu data mining. Strategi ini menggunakan algoritma apriori data mining bisa memberikan sistem strategi promosi penjualan jenis-jenis *ice crime* yang dibahas dalam penelitian.
6. (Putri Mai Sarah Tarigan, 2022), Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang (Studi Kasus : Toko Sinar Harahap). Penelitian ini menjelaskan persediaan barang yang yang tidak dilakukan secara maksimal membuat pihak toko kesulitan untuk menyetok barang yang banyak dibeli

oleh pembeli sehingga menyebabkan ketidak adaan barang di dalam gudang. penulis bertujuan melihat kebutuhan-kebutuhan apasaja yang diperlukan oleh masyarakat sehingga pemilik toko bisa antisipasi persediaan barang di gudang dan tidak mengalami kehabisan.

7. (Debora Kesia Batubara, 2020), Penerapan Algoritma Apriori Untuk Menentukan Persediaan Peralatan dan Mesin. Penelitian ini menjelaskan mencari informasi baru dan dataset dari persediaan alat dan mesin yang ada penulis bisa melihat bagaimana peralatan dan mesin ini selalu ada untuk memenuhi kebutuhan perlengkapan setiap tahunnya agar dana yang dianggarkan tepat untuk kebutuhan-kebutuhan badan pengelolaan aset daerah (BPAD).

2.8 Kerangka Pemikiran

Penjelasan dari kerangka pemikiran ini menggambarkan proses input maksudnya yaitu langkah awal dari data mana yang akan dihitung data disini tidak semua yang dihitung melainkan sampel saja lalu itu di proses menggunakan algoritma apriori seperti menghitung nilai *support* dan *confidencenya* detelah dapat hasil perhitungan manual di lanjutkan dengan proses menggunakan *software* RapidMiner apakah hasil manual dan *softwarena* sama kalau sama akan menghasilkan pengetahuan baru berupa hasil *best seller* penjualan durian di Toko KimHui Batam.



Gambar 2.5 Kerangka Pemikiran Penelitian