

**IMPLEMENTASI *DATA MINING* PADA  
PENGATURAN DATA INVOICE DISTRIBUTOR  
MENGUNAKAN ALGORITMA *FP GROWTH***

**SKRIPSI**



**DISUSUN OLEH :**

**Piter Prasetyo Sudarto**

**190210068**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK DAN KOMPUTER**

**UNIVERSITAS PUTERA BATAM**

**TAHUN 2023**

**IMPLEMENTASI *DATA MINING* PADA  
PENGATURAN DATA INVOICE DISTRIBUTOR  
MENGUNAKAN ALGORITMA *FP GROWTH***

**SKRIPSI**

**Untuk memenuhi salah satu syarat  
memperoleh gelar Sarjana**



**DISUSUN OLEH :**

**Piter Prasetyo Sudarto**

**190210068**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK DAN KOMPUTER**

**UNIVERSITAS PUTERA BATAM**

**TAHUN 2023**

## SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Piter Prasetyo Sudarto  
NPM : 190210068  
Fakultas : Teknik dan Komputer  
Program Studi : Teknik Informatika

Menyatakan bahwa “**Skripsi**” yang saya buat dengan judul:

### **IMPLEMENTASI DATA MINING PADA PENGATURAN DATA INVOICE DISTRIBUTOR MENGGUNAKAN ALGORITMA FP GROWTH**

Adalah hasil karya sendiri dan bukan “duplikasi” dari karya orang lain. Sepengetahuan saya, di dalam naskah Skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka. Apabila ternyata di dalam naskah Skripsi ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah Skripsi ini digugurkan dan gelar akademik yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun

Batam, 29 Juli 2023



Piter Prasetyo Sudarto  
190210068

**IMPLEMENTASI *DATA MINING* PADA PENGATURAN DATA INVOICE  
DISTRIBUTOR MENGGUNAKAN ALGORITMA *FP GROWTH***

**SKRIPSI**

**Untuk memenuhi salah satu syarat  
memperoleh gelar Sarjana**

**Oleh  
Piter Prasetyo Sudarto  
190210068**

**Telah disetujui oleh Pembimbing pada tanggal  
seperti tertera dibawah ini**

**Batam, 27 Juli 2023**



**Koko Handoko, S.Kom., M.Kom.**

**Pembimbing**

## ABSTRAK

Penelitian tentang pemrosesan atau pengolahan data transaksi penjualan dengan menggunakan algoritma *FP Growth* untuk mendukung pengambilan keputusan yang lebih tepat ini dilakukan di PT. Wyssa Artha Sejahtera, sebuah perusahaan distributor. Perusahaan tersebut memiliki banyak data transaksi penjualan setiap harinya, yang terdiri dari berbagai jenis produk, seperti produk minuman dan makanan. Oleh karena itu, penting bagi perusahaan untuk dapat mengekstraksi informasi yang berguna dari data transaksi penjualan tersebut, seperti tren penjualan, produk yang paling diminati oleh pelanggan, dan lain sebagainya. Dalam penelitian ini, metode dokumentasi langsung dengan pihak perusahaan digunakan untuk mengumpulkan data transaksi penjualan dari Januari 2022 hingga Desember 2022. Selain itu, penelitian ini juga menggunakan perangkat lunak *RapidMiner* untuk proses pengolahan data transaksi penjualan. *RapidMiner 9.10* merupakan aplikasi yang sangat tepat untuk menggunakan metode algoritma *FP Growth*, terutama untuk jenis produk makanan dan minuman dimana jumlah data transaksi penjualan semakin banyak. Dalam pengujian yang dilakukan, *RapidMiner 9.10* mampu menemukan set *item* 1, 2, 3, dan 4 hingga mencari *support* dan *Confidence*. Dengan informasi yang dihasilkan dari analisis data transaksi penjualan, perusahaan dapat menentukan strategi yang lebih efektif untuk meningkatkan penjualan dan mendapatkan pola penjualan produk yang memiliki itemset tertentu. Selain itu, hasil penelitian ini juga dapat digunakan sebagai acuan bagi perusahaan-perusahaan lain dalam industri yang serupa untuk meningkatkan efektivitas pengambilan keputusan. Peneliti melakukan analisis dengan menggunakan algoritma *FP Growth*. Algoritma ini digunakan untuk mengekstraksi informasi seperti *support*, *confident*, dan nilai *frequency item-set*. Hasil analisis menunjukkan bahwa *item set* 2, 3, dan 4 memiliki minimal *support* sebesar 20% dan minimal *Confidence* 70%. Setelah data transaksi penjualan terkumpul, Hal ini menunjukkan bahwa produk penjualan di PT. Wyssa Artha Sejahtera produk tertentu seperti Layer Cake Chocolate dan Layer Cake Pandan memiliki penjualan yang lebih baik dibandingkan produk-produk lainnya dalam penentuan nilai *support*

**Kata kunci :** *Association rule, FP-Growth, Aturan asosisasi, RapidMiner*

## **ABSTRACT**

*A research was conducted at PT. Wyssa Artha Sejahtera a distribution company, to explore the processing of sales transaction data using the FP Growth algorithm to support better decision-making. The company deals with a large volume of daily sales transaction data involving various types of products, both wet and dry. The study aimed to extract useful information from the sales data, such as sales trends and the most popular products among customers. Data was collected through observations and direct interviews with the company owner from January 2022 to December 2022, using the proactive method for more comprehensive and accurate information. Upon collecting the sales transaction data, the research identified certain products with higher sales than others. RapidMiner software was utilized for processing the sales data, which proved to be suitable for implementing the FP Growth algorithm, especially for wet and dry product types with increasing sales transaction data. In the testing phase, RapidMiner successfully discovered item sets 1, 2, 3, and 4, along with their corresponding support and confidence values. The results of this research carry significant implications for PT. WYSSA ARTHA SEJAHTERA in making more informed decisions. By analyzing the sales transaction data, the company can devise more effective strategies to boost sales and meet customer demands. Furthermore, the findings can serve as a valuable reference for other companies in similar industries, helping to enhance their decision-making processes. The FP Growth algorithm analysis revealed that item sets 2, 3, and 4 had a minimum support of 20% and a minimum confidence of 70%.*

**Keywords :** *Association, Data Mining ,FP-Growth ,RapidMiner*

## **KATA PENGANTAR**

Terimakasih kepada Tuhan Yang Maha Esa yang telah melimpahkan segala rahmat dan karuniaNya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada Program Studi Teknik Informatika Universitas Putera Batam.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati. Dengan segala keterbatasan, penulis menyadari pula bahwa skripsi ini takkan terwujud tanpa bantuan, bimbingan, dan dorongan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

1. Rektor Universitas Putera Batam: Dr. Nur Elfi Husda, S.Kom., M.SI.
2. Dekan Fakultas Teknik dan Komputer: Welly Sugianto, S.T., M.M.
3. Ketua Program Studi Teknik Informatika: Andi Maslan, S.T., M.SI.
4. Koko Handoko, S.Kom., M.Kom. selaku pembimbing Skripsi pada Program Studi Teknik Informatika Universitas Putera Batam
5. Dosen dan Staff Universitas Putera Batam.

Semoga Tuhan Yang Maha Esa membalas kebaikan dan selalu mencurahkan hidayah serta taufik-Nya, Amin.

Batam, 31 April 2023

Penulis     ,Piter     Prasetyo

Sudarto

## DAFTAR ISI

|  |             |
|--|-------------|
| <b>HALAMAN SAMPUL</b> .....                            | <b>i</b>    |
| <b>HALAMAN JUDUL</b> .....                             | <b>ii</b>   |
| <b>SURAT PERNYATAAN ORISINALITAS</b> .....             | <b>iii</b>  |
| <b>HALAMAN PENGESAHAN</b> .....                        | <b>iv</b>   |
| <b>ABSTRAK</b> .....                                   | <b>i</b>    |
| <b>ABSTRACT</b> .....                                  | <b>ii</b>   |
| <b>KATA PENGANTAR</b> .....                            | <b>iii</b>  |
| <b>DAFTAR ISI</b> .....                                | <b>iv</b>   |
| <b>DAFTAR GAMBAR</b> .....                             | <b>vi</b>   |
| <b>DAFTAR TABEL</b> .....                              | <b>viii</b> |
| <b>BAB 1 PENDAHULUAN</b> .....                         | <b>1</b>    |
| 1.1 Latar Belakang Masalah .....                       | 1           |
| 1.2 Identifikasi Masalah .....                         | 3           |
| 1.3 Batasan Masalah.....                               | 4           |
| 1.4 Rumusan Masalah .....                              | 4           |
| 1.5 Tujuan Penelitian.....                             | 5           |
| 1.6 Manfaat Penelitian.....                            | 5           |
| 1.6.1Manfaat Teoritis .....                            | 5           |
| 1.6.2Manfaat Praktis.....                              | 6           |
| <b>BAB 2 TINJAUAN PUSTAKA</b> .....                    | <b>7</b>    |
| 2.1 Teori Dasar .....                                  | 7           |
| 2.2 Knowledge Discovery in <i>Database</i> (KDD) ..... | 7           |
| 2.3 <i>Data mining</i> .....                           | 9           |
| 2.2.1Langkah-Langkah <i>Data mining</i> .....          | 11          |
| 2.4 Metode <i>Data mining</i> .....                    | 12          |
| 2.4.1 <i>Association rule</i> .....                    | 14          |
| 2.5 Objek Penelitian .....                             | 19          |
| 2.6 <i>Software RapidMiner</i> .....                   | 20          |
| 2.7 Penelitian Terdahulu .....                         | 22          |
| 2.8 Kerangka Pikiran.....                              | 24          |
| <b>BAB 3 METODE PENELITIAN</b> .....                   | <b>25</b>   |
| 3.1 Desain Penelitian .....                            | 25          |
| 3.2. Teknik Pengumpulan Data .....                     | 27          |
| 3.2.1Wawancara .....                                   | 27          |
| 3.2.2Observasi .....                                   | 27          |
| 3.3 Operasional Variabel.....                          | 28          |
| 3.4 Algoritma <i>FP-Growth (Data mining)</i> .....     | 29          |
| 3.5 Lokasi dan Jadwal Penelitian .....                 | 31          |
| 3.5.1Lokasi Penelitian .....                           | 31          |
| 3.5.2Jadwal Penelitian .....                           | 31          |
| <b>BAB 4 PEMBAHASAN</b> .....                          | <b>33</b>   |

|   |           |
|---|-----------|
| 4.1 Hasil analisa dengan algoritma <i>FP-Growth</i> .....                           | 33        |
| 4.2 Proses Implementasi <i>FP-Growth</i> Pada Data Transaksi .....                  | 34        |
| 4.2.1Pengambilan Data.....  | 34        |
| 4.2.2Menyiapkan Dataset .....   | 35        |
| 4.2.3Pencarian Frekuensi <i>Item</i> .....  | 37        |
| 4.2.4Dataset diurutkan berdasarkan <i>Priority</i> .....                            | 37        |
| 4.2.5Memberikan Inisial <i>Itemset</i> .....  | 39        |
| 4.2.6Memberikan <i>Suffix</i> dan <i>Prefix</i> pada <i>FP-Tree</i> .....           | 40        |
| 4.2.7Membangun <i>Conditional Pattern Base</i> dan <i>Conditional FP-Tree</i> ..... | 41        |
| 4.2.8Membangun <i>Frequent Itemset Pattern Generated</i> .....                      | 43        |
| 4.3 Hasil Implementasi <i>Association rule</i> .....                                | 44        |
| 4.4 Pengujian menggunakan <i>RapidMiner</i> .....                                   | 46        |
| 4.5 Hasil Pengujian .....   | 49        |
| <b>BAB 5 PENUTUP</b> .....  | <b>52</b> |
| 5.1 Kesimpulan.....   | 52        |
| 5.2 Saran.....  | 53        |
| <b>DAFTAR PUSTAKA</b>   |           |
| <b>Lampiran</b>   |           |

## DAFTAR GAMBAR

|   |    |
|---|----|
| Gambar 2.1 KDD.....                                 | 8  |
| Gambar 2.2 Tipe <i>Data Mining</i> .....            | 9  |
| Gambar 2.3 Langkah KDD.....                         | 10 |
| Gambar 2.4 Langkah – langkah <i>FP Growth</i> ..... | 16 |
| Gambar 2.5 <i>Association rule</i> .....            | 17 |
| Gambar 2.6 Logo Wyssa .....                         | 19 |
| Gambar 2.7 <i>RapidMiner</i> .....                  | 20 |
| Gambar 2.8 <i>UI RapidMiner</i> .....               | 21 |
| Gambar 3.1 Desain Penelitian .....                  | 25 |
| Gambar 3.2 Bentuk <i>FP-Tree</i> .....              | 29 |
| Gambar 3.3 Lokasi Penelitian .....                  | 31 |
| Gambar 4.1 <i>FP-Tree</i> .....                     | 40 |
| Gambar 4.2 <i>Suffix dan Prefix</i> .....           | 41 |
| Gambar 4.3 <i>Tabulasi</i> .....                    | 47 |
| Gambar 4.4 Import Data.....                         | 48 |
| Gambar 4.5 Data Transformasi .....                  | 48 |
| Gambar 4.6 <i>Modelling</i> .....                   | 49 |
| Gambar 4.7 Hasil Data .....                         | 50 |
| Gambar 4.8 Hasil <i>Graph</i> .....                 | 50 |
| Gambar 4.9 Hasil <i>Description</i> .....           | 51 |
| Lampiran 1 .....                                    | 56 |
| Lampiran 2 .....                                    | 57 |
| Lampiran 3 .....                                    | 57 |
| Lampiran 4 .....                                    | 58 |
| Lampiran 5 .....                                    | 58 |
| Lampiran 6 .....                                    | 59 |
| Lampiran 7 .....                                    | 59 |
| Lampiran 8 .....                                    | 60 |
| Lampiran 9 .....                                    | 60 |
| Lampiran 10 .....                                   | 61 |
| Lampiran 11 .....                                   | 62 |
| Lampiran 12 .....                                   | 62 |
| Lampiran 13 .....                                   | 63 |
| Lampiran 14 .....                                   | 63 |
| Lampiran 15 .....                                   | 64 |
| Lampiran 16 .....                                   | 64 |
| Lampiran 17 .....                                   | 65 |
| Lampiran 18 .....                                   | 65 |
| Lampiran 19 .....                                   | 66 |
| Lampiran 20 .....                                   | 66 |

|  |    |
|--|----|
| <b>Lampiran Surat Penelitian 1</b> ..... | 71 |
| <b>Lampiran LOA 1</b> .....              | 73 |
| <b>Lampiran Turnitin 1</b> .....         | 74 |

## DAFTAR TABEL

|  |    |
|--|----|
| <b>Table 3.1</b> Jadwal Penelitian.....                                    | 32 |
| <b>Table 4.1</b> Data Transaksi.....                                       | 36 |
| <b>Table 4.2</b> Frekuensi Barang.....                                     | 37 |
| <b>Table 4.3</b> <i>Priority Item</i> .....                                | 38 |
| <b>Table 4.4</b> <i>Sortir Item</i> .....                                  | 39 |
| <b>Table 4.5</b> Insial Barang.....  | 40 |
| <b>Table 4.6</b> <i>Conditional Pattern Base</i> .....                     | 42 |
| <b>Table 4.7</b> <i>Conditional FP-Tree</i> .....                          | 43 |
| <b>Table 4.8</b> <i>Frequent Itemset Pattern Generated 2 Itemset</i> ..... | 44 |

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Di era komputisasi ini, kebutuhan manusia akan informasi mendorong pesatnya perkembangan teknologi di bidang TIK. Teknologi yang berkembang didukung oleh sarana dan kebutuhan masyarakat yang beroperasi dengan globalisasi. Sistem informasi dan teknologi informasi memberikan dukungan untuk membuat keputusan yang tepat berdasarkan informasi yang tersedia. Dalam perkembangan teknologi informasi yang ada dapat dengan mudah mengolah data, dapat menghasilkan informasi yang dibutuhkan secara akurat dan tepat waktu, serta biaya yang lebih efektif.

PT. Wyssa Artha Sejahtera merupakan distributor yang bergerak pada bidang pangan yang mendistribusi berbagai merek *Oriental Food Industries Sdn. Bhd.* Distribusi barang yang dilakukan oleh PT. Wyssa Artha Sejahtera merupakan distributor pertama di Batam yang melakukan distribusi merek *Oriental* di area Batam, seperti total jumlah jenis barang yang mencapai 41 tipe barang sehingga distribusi yang dilakukan akan memiliki jumlah transaksi yang besar. Distribusi yang besar ini akan memberikan berbagai macam data terutama *Invoice* yang merupakan dokumen yang berisikan isi transaksi dan pembayaran yang terjadi antara penjual kepada pembeli dalam jumlah yang besar.

*Invoices* yang terdapat pada di PT. Wyssa Artha Sejahtera memiliki berbagai macam data penjualan yang berdampak menumpuknya data-data yang

tidak di olah secara efektif oleh *staff* yang bekerja di PT. Wyssa Artha Sejahtera. Masalah ini menimbulkan banyak kendala terutama pada penjualan yang tidak memiliki peningkatan yang konsisten dan kurangnya efektifitas penjualan dikarenakan pola transaksi tidak pernah dianalisa pola transaksi penjualan tersebut.

*Data mining* adalah konsep yang mencakup kegiatan pengumpulan dan analisis data dengan tujuan menemukan pola yang terkait dalam *volume* data besar (Simanjuntak et al., 2022). Untuk melakukan analisis pola distribusi barang, digunakan teknik *data mining Frequent Pattern Growth* yang juga dikenal sebagai *FP-Growth*, merupakan pengembangan dari algoritma Apriori. Dalam algoritma *FP-Growth*, kekurangan yang ada dalam algoritma Apriori telah diperbaiki. *FP-Growth* adalah salah satu algoritma alternatif yang dapat digunakan untuk menemukan himpunan data yang sering muncul (*Frequent itemset*) dalam sekumpulan data (Fajrin & Handoko, 2018). Metode *FP-Growth* ini didesain untuk melakukan perhitungan berdasarkan data konsultan pengguna dan mengubahnya menjadi data pengetahuan yang akan digunakan sebagai *input* bagi sistem yang dirancang. Hasil dari perhitungan menggunakan algoritma. Penelitian ini memanfaatkan beberapa metode pengumpulan data yang akan dijabarkan pada pembahasan, termasuk wawancara dengan karyawan PT.Wyssa Artha Sejahtera dan pengambilan data yang menyangkut paket yang terkait. Maka teknik-teknik pengumpulan data tersebut dapat menyelesaikan masalah dan memperoleh data yang diperlukan dalam penelitian sebagai berikut, pertama dilakukan wawancara dengan karyawan dan pemilik PT.Wyssa Artha Sejahtera dengan melakukan tanya

jawab untuk memperoleh informasi terkait dengan topik penelitian. Kedua, pengambilan data dilakukan dengan fokus pada *invoices* transaksi penjualan yang terkait dengan topik penelitian. Data yang diperoleh dari kedua metode tersebut akan digunakan untuk menganalisis masalah dan mendapatkan hasil yang akurat dalam penelitian ini.

Dengan mempertimbangkan masalah di atas, penulis akan menerapkan teknik *data mining FP Growth* untuk mengatasinya. Teknik ini memerlukan pengolahan data historis transaksi yang besar pada PT. Wyssa Artha Sejahtera dengan menggunakan pendekatan *tree-based* untuk mencari frequent *itemset*, sehingga lebih efektif daripada teknik Apriori. Dengan menggunakan teknik *FP Growth*, pemetaan *FP-Tree* menjadi semakin efektif ketika terdapat banyak data transaksi dengan *item* yang sama. Hal ini memungkinkan proses pencarian frequent *itemset* menjadi cepat dan efisien.

## **1.2 Identifikasi Masalah**

Berdasarkan latar belakang diatas, maka dapat diidentifikasi masalah-masalah yang timbul, antara lain:

1. Pola penjualan yang ada pada perusahaan PT. Wyssa Artha Sejahtera belum pernah dilakukan perhitungan pencarian pola penjualan sehingga penjualan belum mencapai efektifitas yang optimal.
2. Data yang di tampung terdapat berbagai variabel yang tidak dibutuhkan sehingga perlu di lakukan filter data.
3. Stok yang ada pada gudang tidak sesuai dengan hasil penjualan yang terjadi.

### 1.3 Batasan Masalah

Untuk pembahasan penelitian ini dapat diberikan batasan masalah sebagai berikut:

1. Sumber data yang diteliti merupakan periode Januari 2022 – Desember 2022 pada PT. Wyssa Artha Sejahtera.
2. Metode penelitian *data mining* yang dilakukan untuk *data mining* adalah *Association rule*.
3. Pengolahan data pembantu untuk mendapatkan hasil yang optimal berupa *RapidMiner 9.10*.
4. Algoritma *data mining* yang digunakan merupakan algoritma *FP-Growth*
5. Atribut yang di ambil pada data penelitian nama barang dan no *invoices* yang ada pada transaksi

### 1.4 Rumusan Masalah

Berdasarkan latar belakang diatas, maka dapat dirumuskan masalah-masalah yang timbul, antara lain:

1. Bagaimana rumus perhitungan manual yang akan di implementasikan pada *FP-Growth* ?
2. Bagaimana implementasi metode *data mining FP-Growth* untuk mengatasi masalah yang ada pada PT. Wyssa Artha Sejahtera ?
3. Bagaimana aplikasi pembantu *RapidMiner* dapat membandingkan hasil dari pengolahan data ?

## 1.5 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka dapat diketahui tujuan dari penelitian, sebagai berikut:

1. Untuk perhitungan manual akan menggunakan formula antara nilai *confident* dan nilai *support* yang muncul dari suatu *itemset* kemudian akan mendapatkan hasil dari perhitungan manual kemudian hasil tersebut akan dimasukkan pada *Frequent Pattern Tree*.
2. Untuk implementasi penyelesaian masalah pada PT. Wyssa Artha Sejahtera akan digunakan algoritma *FP-Growth* untuk menentukan pola penjualan.
3. Untuk Aplikasi pembantu *RapidMiner* membandingkan hasil dari perhitungan dengan rumus otomatis yang tersedia kemudian menjadi alat perbandingan.

## 1.6 Manfaat Penelitian

Berdasarkan latar belakang di atas jika dilakukan penelitian akan memberikan beberapa dampak positif, diantaranya:

### 1.6.1 Manfaat Teoritis

1. Penulis dapat mengidentifikasi masalah-masalah yang dihadapi oleh PT. Wyssa Artha Sejahtera dengan metode *data mining* dan memberikan solusi.
2. Menambah pengalaman, pemahaman, ilmu, dan kemampuan peneliti dalam mempelajari metode *data mining*.
3. Peneliti mampu memberikan ilmu *data mining* terhadap permasalahan nyata yang dihadapi perusahaan sehingga dapat diterapkan penggunaannya.

### **1.6.2 Manfaat Praktis**

1. Penulis menerapkan metode *data mining* dalam menentukan pola penjualan yang dilakukan PT.Wyssa Artha Sejahtera sehingga dapat memberikan solusi dalam meningkatkan efektifitas penjualan.
2. Perusahaan UMKM lainnya dapat mengikuti metode ini terhadap sebagai contoh pola transaksi dari perusahaan distributor untuk menentukan pola penjualan pada perusahaan tersebut.
3. Memperoleh ilmu tentang *data mining* yang menggunakan *Frequent Pattern Growth* sehingga PT.Wyssa Artha Sejahtera mampu membeli barang yang tepat sesuai dengan pola penjualannya.

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Teori Dasar**

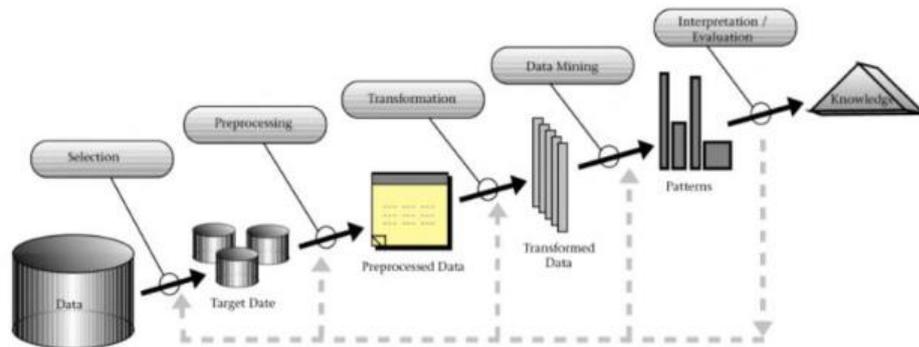
Peneliti akan menjelaskan hal-hal apa saja yang dapat membantu peneliti dalam menyelesaikan topik ini. Diantaranya adalah mengenai *Knowledge Discovery in Database (KDD)*, *Data mining*, algoritma *FP-Growth*, dan juga membahas mengenai *software* pendukung yaitu *RapidMiner* berikut penjelasannya.

##### **2.1.1 Knowledge Discovery in Database (KDD)**

*Knowledge Discovery in Database (KDD)* merupakan pengumpulan dan pemanfaatan data historis dengan tujuan mengidentifikasi keteraturan, pola, atau hubungan dalam catatan dalam *database* besar yang sebelumnya tidak diketahui dan memiliki nilai (Bunda, 2020). Dalam pengumpulan data, KDD menggunakan data historis untuk menemukan pengetahuan, informasi, keteraturan, pola atau hubungan dalam data yang berukuran besar. KDD merupakan suatu kegiatan yang dilakukan dalam rangka memperoleh informasi baru dari data yang ada (Fajrin & Handoko, 2018)

Penemuan pengetahuan dalam *database (KDD)* terdiri dari ekstraksi informasi implisit potensial dan satu set data yang tidak diketahui (Simanjuntak & Elisa, 2019). Dengan memasukkan hasil proses *data mining*, yaitu proses penemuan pengetahuan, informasi diubah menjadi hasil yang akurat dan lebih mudah dipahami. Maka berdasarkan pengertian di atas dapat disimpulkan bahwa pengertian dari KDD adalah proses penting untuk menemukan dan mengenali

pola dalam data dan bentuk yang dihasilkan baru, berguna, dan mudah dipahami. Hal ini dapat diperoleh dari informasi yang dihasilkan oleh *data mining*.



Gambar 2.1 KDD

Adapun penentuan dalam melakukan pemrosesan *data mining* dimulai dari data yang akan diterima, *selection*, *transformation*, *data mining*, *evaluation/interpretation*, sehingga menjadi data yang berguna sebagai berikut (Bunda, 2020):

1. *Selection*, adalah Pemilihan data (seleksi) dari sekumpulan data operasional harus dilakukan sebelum tahap ekstraksi KDD dimulai. Data terpilih yang digunakan dalam proses *data mining* disimpan dalam file terpisah dari *database* produksi.
2. *Preprocessing*, merupakan Langkah proses *data mining*, yang harus menjalankan proses pembersihan pada data yang menjadi bagian awal KDD. Proses pembersihan termasuk menghapus data duplikat, verifikasi data yang tidak konsisten, merubah kesalahan data, dan berbagai contoh lainnya.

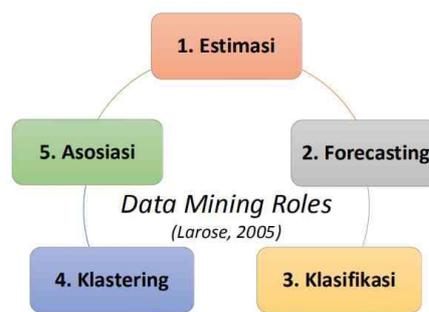
3. Transformasi, proses mentransformasikan data terpilih sehingga data tersebut sesuai untuk proses *data mining*. Proses ini merupakan proses yang memiliki algoritma tertentu dan sangat tergantung pada jenis atau pola informasi yang diambil dalam *database*.

4. *Data mining*, proses menemukan pola atau info yang ada pada data yang dipilih dengan menggunakan suatu teknik serta metode tertentu. Pada bentuk teknik, jenis metode, ataupun langkah algoritma *data mining* memiliki fungsi masing masing. Maka diperlukan pemahaman tentang teknik, algoritma, serta metode yang mendalam agar dapat menggunakan dengan efektif.

5. *Interpretation / Evaluation*, pola informasi berupa hasil langkah-langkah *data mining* yang harus diberikan dalam format yang mudah dipahami oleh peneliti demi kepentingan pembaca. Fase ini merupakan bagian dari proses KDD yang disebut interpretasi. Fase ini memeriksa apakah pola atau bentuk informasi yang tidak sesuai dengan fakta atau hipotesis yang ada.

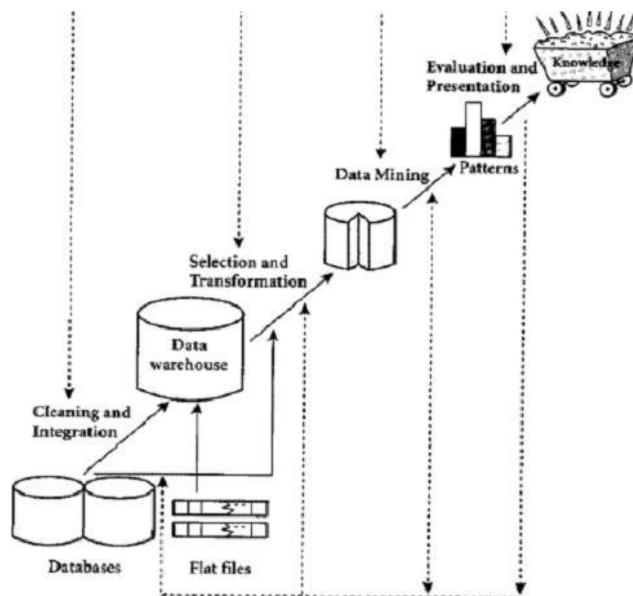
### 2.1.2 Data mining

*Data mining* adalah dapat bermanfaat dan dimengerti dalam suatu *database*



Gambar 2.2 Tipe Data mining

yang sangat besar, *data mining* merupakan proses *iterative* dan *interactive* untuk menemukan pola atau model baru yang sempurna mendefinisikan bahwa pengambilan keputusan di waktu yang akan datang dapat dibantu dengan *data mining* yang berisi pencarian trend atau pola yang diinginkan dalam *database* besar, yang mungkin saja menggunakan perangkat pendukung keputusan yang lain dari pola-pola yang dikenali perangkat tertentu yang dapat memberikan suatu analisa data yang berguna dan berwawasan untuk kemudian dapat dipelajari serta dipahami dengan lebih teliti.



Gambar 2.3 Langkah *Data mining*

Adapun dua pernyataan yang menyatakan manfaat dari *data mining*, sebagai berikut:

1. Karena fakta bahwa pola data potensial seringkali tidak terlihat, menemukan pola data dapat membuat frustrasi bagi pembuat keputusan yang tidak berpengalaman.

2. Berdasarkan pemahaman di atas bahwa jumlah data terlalu besar untuk dianalisis secara manual, disimpulkan bahwa *data mining* adalah pencarian untuk membuat keputusan di masa depan dengan menggunakan pola yang diinginkan dalam *database* yang besar. Dengan menggunakan teknik khusus dari kumpulan data besar, penambahan data dapat mengekstrak informasi dan wawasan penting.

### **2.1.3 Langkah-Langkah *Data mining***

Berdasarkan pengertian menurut (Rahmawati & Merlina, 2018) dalam melakukan *data mining* diperlukan langkah-langkah yang tepat sebagai berikut :

1. *precise statement of the problem* yakni sebelum menggunakan perangkat lunak pengolah *data mining*. Masalah yang ditemukan oleh analis harus jelas tentang pertanyaan tersebut. Diperlukan bentuk penjelasan yang tepat untuk masalah yang sedang ditemui saat membuat solusi.
2. Poin utama dari fase ini adalah identifikasi dan eliminasi (pembersihan) data yang salah kode, transformasi data, pemilihan subset dataset, dataset, langkah pertama dalam seleksi, deskripsi data dan visualisasi.
3. *Model building and validation* Pada fase ini, kinerja prediktif terbaik ditentukan dan jenis pemodelan dipertimbangkan.
4. *Deployment*. Pada langkah fase Pada fase ini, kita membuat prediksi (generasi). Artinya, pilih aplikasi yang sesuai dengan pemodelan. Kemudian mengimplementasi data yang akan diolah. Penentuan nilai *support*, *confident*, dan *frequent item-set* berdasarkan aturan asosiatif. Berikut penjelasannya :

1. Nilai *Support* (*s*) merupakan sebuah nilai persentase kemungkinan gabungan jumlah *item* yang berbeda.

$$\text{Support}, s(X \rightarrow Y) = \frac{(X \cup Y)}{N}$$

Yang dimana  $X \cup Y$  adalah sebuah jumlah gabungan transaksi yang terdapat pada *item* X dan Y, kemudian N adalah jumlah total yang ada pada seluruh transaksi penjualan barang.

2. Nilai *confident* (*c*) adalah suatu jumlah presentase ketepatan pada aturan asosiasi yang dihitung.

$$\text{Confident}, c(X \rightarrow Y) = \frac{(X \cup Y)}{X}$$

Yang dimana  $X \cup Y$  merupakan total penjualan gabungan *item* X dan Y. Sementara nilai X berupa jumlah penjualan yang *item* X.

3. *Frequent item-set*, tahap pertama terdiri dari menentukan set *item* yang sering. Sering di sini berarti kombinasi set *item* yang sering terjadi dalam suatu dataset (data transaksional). Aturan yang dibuat kemudian untuk menghasilkan tingkat kepercayaan yang tinggi. Dengan melakukan kemungkinan kombinasi set *item* berdasarkan *k-itemset* (penghasilan set *item* yang sering), di mana k mengacu pada jumlah *item* yang akan digabungkan.

#### 2.1.4 Metode *Data mining*

Pada penelitian ini penulis menggunakan metode *data mining* yang bernama aturan asosiasi. Istilah aturan asosiasi adalah teknik *data mining* yang kuat untuk menemukan korelasi dan hubungan antara objek dalam *database*, didasarkan pada

analisis statistik dan kecerdasan buatan. Teknik ini sangat tepat untuk menganalisis korelasi antar objek, karena mempertimbangkan interaksi kondisional di antara kumpulan data *input*, dan menghasilkan aturan keputusan dalam bentuk *IF-THEN*(Ait-Mlouk et al., 2017).

*Data mining* pertama kali dilakukan oleh Charu C. Aggarwal untuk pertama kalinya, menganalisis *database* transaksional. Aturan asosiasi biasanya didefinisikan sebagai implikasi dari bentuk:  $A \rightarrow B$  seperti  $A, B \subset I$  dan  $A \cap B = \emptyset$ . Setiap aturan terdiri dari dua set *item* A dan B yang berbeda, di mana A disebut mendahului dan B disebut konsekuensial. Misalnya  $\{\text{Pengemudi}\} \rightarrow \{\text{Kendaraan}\}$ , menunjukkan bahwa ada hubungan yang kuat antara dua *item*  $\{\text{Pengemudi}, \text{Kendaraan}\}$ .

Menurut (Antari et al., 2016) menyatakan bahwa menemukan aturan asosiasi antar kombinasi *item* merupakan teknik *data mining*. Hasil pengolahan data dan perhitungan tertentu kepentingan (*action of interest*) adalah untuk menentukan aturan asosiasi. Secara umum, aturan asosiasi memiliki dua indikator penting. Sebagai berikut :

1. *Support*, jumlah semua transaksi dari probabilitas bahwa konsumen membeli beberapa produk secara bersamaan. Dari semua transaksi yang tersedia, metrik ini menentukan apakah suatu *item*/kelompok *item* layak untuk diteliti karena nilai *Confidence*-nya. Besarnya dominasi menunjukkan bahwa barang X dan Y dibeli secara bersamaan.
2. *Confidence* (tingkat kepercayaan) adalah produk yang dibeli secara bersamaan dari beberapa kemungkinan kejadian dimana produk

tersebut pasti akan dibeli (Contohnya :frekuensi penjualan *item Y* dibeli apabila konsumen membeli *item X*)

Dua parameter (*Support* dan *Confidence*) digunakan untuk menentukan aturan ditentukan oleh pengguna untuk dibandingkan sebagai batas akhir (*Threshold*). Batasan tersebut merupakan batas minimal dari nilai *support* yang merupakan *level* minimum *support* dan batas minimal nilai *Confidence* adalah minimal *Confidence*. Berikut ini penjelasan masing – masing metode *data mining* :

#### **2.1.4.1 Association rule**

Tujuannya adalah untuk menunjukkan nilai asosiatif antara jenis produk yang dibeli pelanggan dan memungkinkan mereka untuk melihat pola dalam bentuk produk yang sering dibeli pelanggan. Dengan mengetahui jenis produk yang biasa dibeli, hal ini dapat dijadikan dasar untuk memutuskan produk mana yang pantas untuk diiklankan kepada pelanggan tersebut.

*Association rule* adalah konsep penting dalam analisis data yang digunakan untuk mengidentifikasi hubungan antara *item-item* dalam *dataset*. Dalam penelitian yang dilakukan oleh Agrawal, Imielinski, dan Swami (1993), *association rule* dijelaskan sebagai aturan yang menghubungkan antara suatu kelompok *item* atau atribut dengan kelompok *item* atau atribut lainnya dalam dataset. Aturan asosiasi ini sering digunakan dalam *data mining* untuk menemukan keterkaitan antara *item* atau atribut yang sering muncul bersama dalam suatu konteks.

Proses penemuan *association rule* dilakukan melalui dua langkah utama, yaitu pembentukan kandidat *rule* dan pengujian kandidat *rule*. Dalam langkah

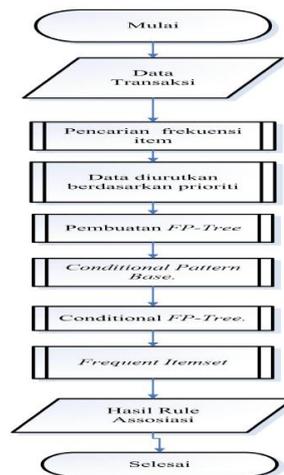
pertama, kandidat *rule* dibangun dengan memilih *item* atau atribut yang sering muncul bersama dalam dataset. Kemudian, pada langkah kedua, kandidat *rule* tersebut diuji dengan menghitung tingkat dukungan dan tingkat kepercayaan untuk memilih *rule* yang memiliki nilai di atas batas ambang tertentu. Tingkat dukungan mengukur frekuensi kemunculan *rule* dalam dataset, sedangkan tingkat kepercayaan mengukur sejauh mana *rule* tersebut dapat dipercaya.

Penerapan *association rule* dalam analisis data memiliki banyak manfaat dalam berbagai bidang. Misalnya, dalam analisis keranjang belanja, *association rule* dapat membantu toko atau perusahaan untuk mengidentifikasi *item* yang sering dibeli bersama-sama oleh konsumen. Dengan mengetahui hubungan antara *item*, toko dapat meningkatkan strategi penjualan lintas produk dan meningkatkan pengalaman belanja pelanggan. Selain itu, *association rule* juga digunakan dalam rekomendasi sistem untuk menghasilkan rekomendasi produk yang relevan berdasarkan perilaku konsumen dan preferensi mereka.

#### **2.1.4.2 Algoritma *FP-Growth***

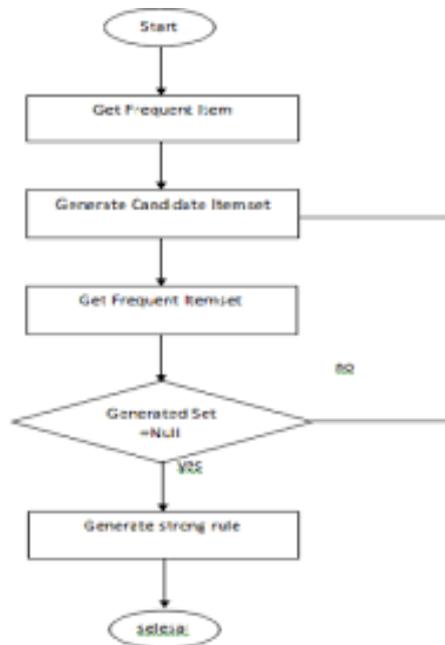
*FP-Growth* memiliki keunggulan dalam memproses data yang besar dan kompleks. Hal ini dikarenakan algoritma ini hanya memerlukan satu pemindaian data transaksi untuk membangun *FP-Tree*, yang dapat menghemat waktu pemrosesan data. Selain itu, *FP-Growth* juga memiliki kemampuan untuk menghasilkan *itemset* yang lebih besar dan kompleks dibandingkan dengan algoritma-aplikasi lainnya (Kumar & Mohbey, 2022)

Dalam konteks penelitian, *FP-Growth* dapat digunakan untuk mengatasi masalah yang melibatkan analisis pola distribusi barang dalam transaksi. Dengan menerapkan *FP-Growth* pada data historis transaksi yang besar, peneliti dapat mengidentifikasi pola-pola frekuensi tinggi yang dapat memberikan wawasan tentang tren penjualan, asosiasi antara *item* barang, dan preferensi pelanggan. Dengan demikian, *FP-Growth* memberikan kontribusi dalam pengambilan keputusan yang lebih baik dalam hal strategi pemasaran, manajemen persediaan, dan pengembangan produk.(Bunda, 2020).



Gambar 2.4 Langkah – langkah FP Growth

*FP-Growth* menggunakan pendekatan berbasis pohon untuk menghasilkan



Gambar 2.5 Association rule

pola-pola yang sering muncul dalam data transaksi. Algoritma ini bekerja dengan membangun *FP-Tree* dari data transaksi yang ada dan kemudian melakukan pemrosesan berbasis pohon untuk mengidentifikasi pola-pola frekuensi tinggi. *FP-Growth* memiliki keunggulan dibandingkan metode *data mining* lainnya, seperti Apriori, karena tidak memerlukan proses pemindaian ulang *dataset* setelah membangun *FP-Tree*, sehingga lebih efisien dalam hal kecepatan pemrosesan.

$$\text{Confident } P(A \cap B) = \frac{\Sigma(A \cap B)}{\Sigma A} \times 100\%$$

Rumus Confident

$$\text{Support } (A) = \frac{\text{Jumlah Transaksi Yang Mengandung } A}{\text{Total Transaksi}} \times 100\%$$

Rumus Support

$$\text{Support (A,B)} = \frac{\text{Jumlah Transaksi Yang Mengandung A dan B}}{\text{Total Transaksi}} \times 100\%$$

### 2.4.1.3 Algoritma Apriori

Algoritma Apriori merupakan algoritma yang sangat efisien untuk menentukan jumlah set *item* umum (Data et al., 2020). Dasar pemikiran dari algoritma ini adalah jika suatu himpunan *item* merupakan himpunan *item* yang sama, maka semua subset (bagian) dari himpunan *item* tersebut juga sama, dan sebaliknya. Contoh: Misalnya, jika set *item* A tidak umum (tidak sering muncul dalam transaksi), *item* yang digabungkan ke set *item* A tidak akan membuat set *item* A menjadi umum (sering terjadi dalam transaksi). Penggunaan algoritma ini untuk penelitian yang dilakukan dikarenakan pola pembelian *retail* dalam pembelian bahan pokok untuk menjual ke konsumen. Berdasarkan pola penjualannya tiap *retail* membeli barang sesuai dengan tipe *retail* tersebut dan fokusnya tertentu.

Algoritma Apriori bekerja dengan membangun kandidat *itemset* dari dataset transaksi, kemudian melakukan pemindaian ulang dataset untuk menghitung frekuensi kemunculan setiap kandidat. *Itemset* yang memenuhi tingkat dukungan minimum ditetapkan sebagai *itemset* frekuensi tinggi. Proses ini dilakukan secara berulang dengan meningkatkan ukuran *itemset* yang diuji setiap iterasi, hingga tidak ada lagi kandidat yang memenuhi tingkat dukungan minimum.

Penerapan algoritma Apriori dalam *data mining* dapat memberikan wawasan yang berharga dalam berbagai bidang, seperti analisis keranjang belanja, rekomendasi produk, dan pemrosesan teks. Dengan mengidentifikasi asosiasi

antara *item*, Apriori dapat membantu organisasi dalam mengambil keputusan yang lebih baik, seperti menyusun strategi pemasaran yang lebih efektif, meningkatkan penjualan lintas produk, dan memahami perilaku konsumen. Meskipun algoritma Apriori cukup sederhana, namun memiliki keunggulan dalam menemukan asosiasi yang penting dalam dataset transaksi yang besar dan kompleks.

## 2.2 Objek Penelitian

PT. Wyssa Artha Sejahtera, yang didirikan pada tahun 2018 di Batam, merupakan sebuah perusahaan distributor yang memiliki fokus bisnis di sektor pangan, khususnya makanan ringan. Perusahaan ini memainkan peran penting dalam rantai pasok pangan dengan menjalin kerja sama erat dengan produsen makanan, baik di dalam negeri maupun luar negeri. Kolaborasi ini dirancang untuk memastikan kelancaran proses penyaluran produk ke berbagai pasar dan pelanggan.



Gambar 2.6 Logo Wyssa

Salah satu tujuan utama pendirian PT. Wyssa Artha Sejahtera adalah untuk mendukung kebutuhan pangan masyarakat Indonesia. Dengan menyediakan akses yang lebih mudah terhadap makanan ringan yang berkualitas, perusahaan ini berkontribusi dalam memenuhi kebutuhan gizi dan memberikan variasi pilihan makanan yang lezat kepada konsumen di seluruh Indonesia. Selain itu, PT. Wyssa

Artha Sejahtera juga berkomitmen untuk menciptakan lapangan kerja dan kesempatan berkarir bagi masyarakat, sehingga dapat membantu mengurangi tingkat pengangguran di daerah sekitar dan berkontribusi pada pertumbuhan ekonomi secara keseluruhan.

Dengan fokus pada penyaluran makanan ringan berkualitas dan kerja sama yang solid dengan produsen, PT. Wyssa Artha Sejahtera berupaya menjadi mitra pilihan yang dapat diandalkan dalam industri distribusi pangan. Perusahaan ini terus berinovasi dan meningkatkan layanan untuk memenuhi kebutuhan pasar yang terus berkembang dan memberikan kontribusi positif bagi masyarakat dan perekonomian Indonesia secara keseluruhan.

### **2.3 *Software RapidMiner***

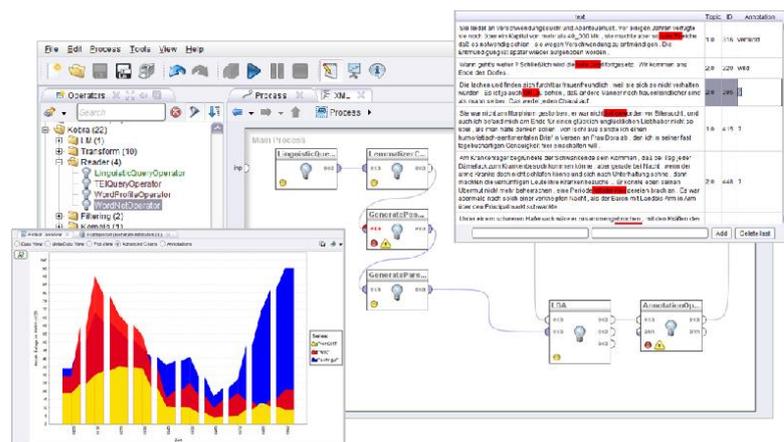
*RapidMiner* adalah sebuah perangkat lunak yang dapat digunakan untuk mengolah data dengan berbagai macam fungsi dan operator unik. Perangkat ini digunakan mulai dari tahap pengolahan dan peletakan data hingga machine learning yang mampu menghasilkan perubahan pada data yang akan dimanipulasi (Fajrin & Handoko, 2018).



Gambar 2.7 *RapidMiner*

*RapidMiner* adalah sebuah platform analitik yang terkemuka di bidang *data mining* dan analisis prediktif. Platform ini menyediakan berbagai fitur yang kuat untuk menggali pengetahuan dari data, mengembangkan model prediktif, dan melakukan analisis statistik. Dalam penggunaannya, *RapidMiner* memiliki potensi yang signifikan untuk memberikan manfaat bagi organisasi dalam berbagai konteks.(Bunda, 2020)

Selain itu, *RapidMiner* juga memiliki kemampuan untuk mempresentasikan hasil analisis dengan cara yang intuitif dan mudah dipahami. Platform ini menyediakan fitur visualisasi data dan laporan yang memungkinkan



Gambar 2.8 UI *RapidMiner*

pengguna untuk secara efektif menyajikan temuan-temuan penting kepada pemangku kepentingan. Keunggulan ini dapat meningkatkan kemampuan pengambilan keputusan yang lebih baik serta memudahkan komunikasi hasil analisis kepada tim atau pihak yang terlibat dalam proses pengambilan keputusan.(AHMAD ADRI, 2021)

## 2.4 Penelitian Terdahulu

Berdasarkan penelitian terdahulu yang sesuai dengan objek yang diteliti serta metode penelitian, berikut ini merupakan penelitian-penelitian yang berkaitan dengan judul yang sedang diteliti :

1. (Hamid Mughal, 2018) “*Data mining has become a significant platform for retrieving valuable information from the World Wide Web. Users find it convenient to upload and download data through web data mining techniques, tools, and algorithms.*” (IJACSA) *International Journal of Advanced Computer Science and Applications*, Vol. 9, No. 6, 2018
2. (Kurnia et al., 2019)“ *This research investigates data mining market basket analysis to understand sales patterns at O! Fish restaurant using the Apriori algorithm. The study presents a website-based application to analyze consumer purchasing patterns, which can then be used to recommend promotion development strategies for the restaurant.*” *International Conference on Advance and Scientific Innovation (ICASI) Conf. Series 1175 (2019) 012047*
3. (Dogan & Birant, 2021)“*The article "Machine learning and data mining in manufacturing" discusses the benefits of employing machine learning methods in the manufacturing sector, provides strategies to tackle specific challenges, and proposes potential avenues for future research exploration.*” *Expert Systems with Applications* 166(2):114060
- 4.(Prasetya et al., 2021)” *Algoritma Apriori memiliki potensi untuk digunakan dalam sistem rekomendasi barang. Pengujian dengan 20 data transaksi dan nilai support minimum 20% serta minimum Confidence 70% menunjukkan hasil yang*

konsisten dengan persyaratan tersebut. Dalam pengujian dengan menggunakan algoritma Apriori dan aplikasi yang dibuat, hasil yang diperoleh sesuai dengan nilai *support* dan *Confidence* yang diharapkan.” Jurnal JTIIK (Jurnal Teknologi Informasi Dan Komunikasi), 5(4), 354–361.

5.(Sibarani, 2020) “Dengan bantuan *data mining*, Yati Cosmetics dapat mengetahui tren penjualan kosmetik yang paling diminati oleh konsumen dan mengoptimalkan persediaan kosmetik yang tersedia. Analisis ini memberikan manfaat dalam meningkatkan sistem penjualan mereka dan memberikan pemahaman yang lebih mendalam tentang preferensi pelanggan. Hasil analisis data juga memberikan aturan-aturan yang membantu dalam pengambilan keputusan, memungkinkan toko untuk meningkatkan efisiensi dan efektivitas dalam bisnis mereka.” Computer and Science Industrial Engineering (COMASIE), 3(3), 102–111.

6.(Rahmawati & Merlina, 2018) “Dari hasil penelitian ini, ditemukan bahwa produk yang paling banyak terjual adalah *drum dc* dan *toner*. Kesimpulan yang dihasilkan menunjukkan bahwa produk toner dan drum dc memiliki tingkat dukungan (*support*) sebesar 50% dan tingkat kepercayaan (*Confidence*) sebesar 66,67%. Selain itu, produk *drum dc* dan *toner* juga memiliki tingkat dukungan sebesar 50% dan tingkat kepercayaan sebesar 85,71%.”PIKSEL : Penelitian Ilmu Komputer Sistem Embedded and Logic p-ISSN: 2303-3304 | e-ISSN: 2620-3553Vol 6 No 1 (2018): Maret 2018.

7.(Simanjuntak & Elisa, 2019) “Kartu perdana prabayar sangat mudah didapatkan dengan harga yang terjangkau dan rata-rata murah di berbagai lokasi. Penelitian

ini bertujuan untuk memilih kartu seluler terbaik di Kota Batam dengan menggunakan *Data mining* dan metode Asosiasi berbasis algoritma apriori. Penelitian ini diimplementasikan menggunakan perangkat lunak Tanagra 4.1.”  
Journal Information System Development (ISD), 4(2)

## 2.5 Kerangka Pikiran

Bentuk rangka pemikiran yang akan dikembangkan pada penelitian ini adalah sebagai berikut:



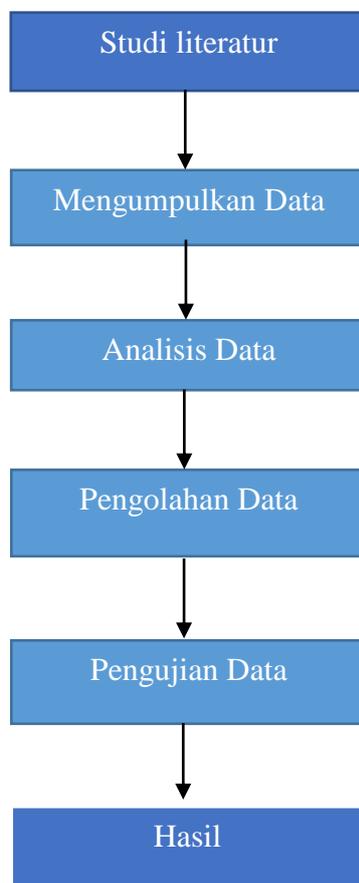
Pada peta di atas, pada proses pertama penelitian ini data tentang penjualan barang sebagai proses *input*. bentuk *input* data merupakan produk/barang seperti merk dan jenis produk, serta jumlah total transaksi yang dilakukan ini serta disaring terlebih dahulu. Selain itu, data ini diproses oleh algoritma setelah data diproses dan mencapai nilai dukungan minimum, apriori kepercayaan minimal. Kemudian uji data yang akan kita gunakan menggunakan perangkat lunak *RapidMiner* untuk mendapatkan hasil yang akurat dan menemukan hasil penelitian tersebut sehingga hasil dari proses adalah produk yang paling sering dibeli konsumen berguna untuk toko yang ingin mengetahui pola penjualan.

## BAB III

### METODE PENELITIAN

#### 3.1 Desain Penelitian

Desain penelitian yang akan dilakukan pada penelitian berikut akan sesuai dengan *flowmap* berikut :



Gambar 3.1 Desain Penelitian

Desain penelitian merupakan langkah penelitian yang dilakukan dalam penelitian ini. Rancangan penelitian di atas menggambarkan proses dari awal

hingga akhir kegiatan. Berikut ini adalah penjelasan langkah desain penelitian di atas, yaitu:

1. Langkah pertama, mengidentifikasi masalah yang sedang diselidiki dan kemudian mengidentifikasi judul yang sedang diselidiki. Fase ini juga mengidentifikasi masalah, yaitu masalah pada PT.Wyssa Artha Sejahtera yang sedang dilakukan pada peneliti. Selanjutnya adalah analisis masalah yang meneliti hal-hal dan hal-hal secara rinci dan merakit dan meneliti bagian-bagian yang dijelaskan dari bagian-bagian.
2. Langkah kedua, data dikumpulkan untuk mencapai tujuan penelitian guna memperoleh informasi yang diperlukan. Dengan mengumpulkan data pada tahun Januari 2022-Desember 2022, peneliti memperoleh data observasi langsung dari pemilik toko, wawancara langsung ke PT Wyssa Artha Sejahtera, dan awalnya berasal dari bahan penelitian literatur tentang *data mining* dan referensi ke jurnal dan buku nasional dan internasional.
3. Langkah ketiga, Data yang dikumpulkan dan dianalisis disiapkan dan diproses dalam proses *data mining*. Yaitu, algoritma *FP-Growth* untuk menentukan nilai *support* dan *Confidence* minimum untuk setiap produk yang dijual.
4. Langkah keempat, Implementasi dalam penelitian ini adalah penerapan data penjualan barang yang telah di filter menggunakan teknik

algoritma *FP-Growth* yang akan di uji dengan perangkat pembantu *RapidMiner*.

5. Langkah kelima, hasil tersedia setelah semua tahapan atau proses penelitian. Kemudian akan menggunakan algoritma *FP-Growth* untuk menguji hasil menggunakan perangkat lunak *RapidMiner* dan memeriksa hasilnya untuk melihat apakah cocok dengan hasil tes.
6. Langkah terakhir yang merupakan tahap akhir penelitian, pada tahap ini akan menemukan hasil penelitian selidiki set *item* terlaris dan tentukan presentase produk penjualan di perusahaan.

### **3.2. Teknik Pengumpulan Data**

Teknik pengumpulan data terhadap penelitian yang sedang teliti memiliki beberapa penjelasan sebagai berikut:

#### **3.2.1 Wawancara**

Wawancara adalah cara peneliti mengolah data tentang informasi dengan bertanya langsung kepada informan. Keberhasilan suatu wawancara ditentukan oleh banyak faktor, antara lain pewawancara, informasi, topik survei kuesioner, dan faktor keadaan di mana wawancara sedang berlangsung (Simanjuntak et al., 2022). Kesimpulan teori di atas, tahap wawancara yang ini dilakukan oleh peneliti secara langsung dengan Kepala Gudang. PT Wyssa Artha Sejahtera untuk mendapatkan data penelitian.

#### **3.2.2 Observasi**

Observasi adalah metode pengumpulan data dengan cara mengamati perilaku dalam situasi tertentu, mencatat secara sistematis peristiwa yang diamati,

dan memahami peristiwa yang diamati (Ashma Nurmeila et al., 2020). Pengamatan dalam penelitian ini terhadap pembelian, persediaan produk, dan data distribusi produk yang berada perusahaan PT. Wyssa Artha Sejahtera.

### 3.3 Operasional Variabel

Variabel perlu didefinisikan secara operasional sehingga hubungan antara satu variabel dengan variabel lainnya dan pengukurannya dapat dengan mudah ditemukan. Tanpa variabel instrumental, peneliti kesulitan memutuskan untuk mengukur hubungan antar variabel, tetapi ini masih konseptual. Dalam penelitian ini, variabel instrumentalnya adalah:

- 1) Data Transaksi

Data transaksi adalah jumlah total transaksi yang terjadi dalam kurun waktu 12 bulan sejak Januari 2022 hingga Desember 2022. Ini dipengaruhi oleh jumlah transaksi untuk setiap produk. Ini akan menentukan kombinasi *item* mana yang paling laku.

- 2) Nama produk

Nama produk umumnya merupakan merek produk yang dibeli dari toko retail dan merupakan produk sebagai pilihan persediaan. Merek merupakan variabel *input* dalam penelitian ini. Merek produk apa yang paling banyak diminta pembeli dan merek apa yang tersedia untuk perusahaan PT. Wyssa Artha Sejahtera.

- 3) Jenis produk

Jenis produk adalah jenis produk yang paling dicari dan dibeli pembeli, dan merupakan jenis produk pilihan inventaris. Penulis

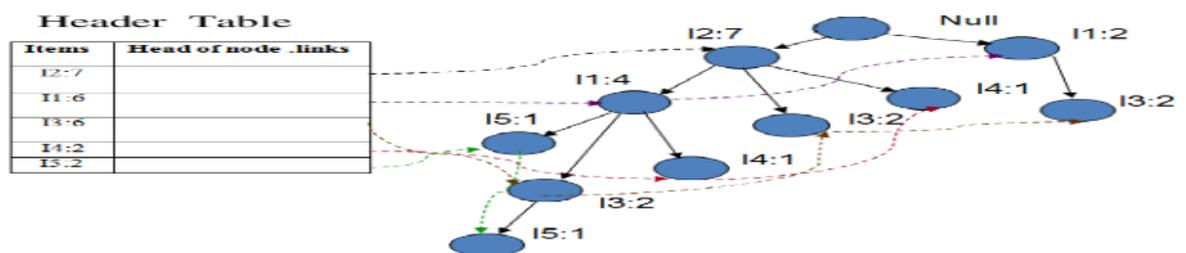
memilih produk kering dan basah sebagai bahan penelitian untuk penelitian ini.

### 3.4 Algoritma *FP-Growth* (Data mining)

Dalam penelitian ini, digunakan metode penelitian yang menggunakan algoritma *FP-Growth* (aturan asosiasi). Metode algoritma *FP-Growth* ini termasuk sebagai salah satu teknik *Market Basket Analysis*. Penelitian ini akan mengidentifikasi produk terlaris yang dijual oleh PT. Wyssa Artha Sejahtera menggunakan tabel transaksi penjualan. Data transaksi tersebut dapat membantu menganalisa pola pembelian yang dilakukan oleh pelanggan dan memungkinkan PT Wyssa Artha Sejahtera sebagai distributor untuk menjual berbagai produk dengan lebih efektif.

**Table 1: A dataset with nine transactions.**

| TID | List of items |
|-----|---------------|
| T1  | I1,I2,I5      |
| T2  | I2,I4         |
| T3  | I2,I3         |
| T4  | I1,I2,I4      |
| T5  | I1,I3         |
| T6  | I2,I3         |
| T7  | I1,I3         |
| T8  | I1,I2,I3,I5   |
| T9  | I1,I2,I3      |



Gambar 3.2 Bentuk *FP-Tree*

Analisis asosiasi merupakan teknik penting dalam *data mining* yang menjadi dasar bagi banyak teknik *data mining* lainnya. Tahapannya meliputi pembuatan algoritma yang efisien, khususnya pada analisis pola frekuensi tinggi. Dalam analisis asosiasi, ada dua parameter penting yaitu dukungan dan keandalan. Dukungan (*support*) adalah persentase kemunculan kombinasi elemen dalam *database*, sedangkan keandalan (*Confidence*) menunjukkan kekuatan hubungan antar elemen. Teknik analisis asosiasi sangat berguna dalam berbagai industri, seperti dalam urusan penjualan ritel. Adapun langkah perlakuan *FP-Growth*

Iterasi 1: Memindai *database* dan menghitung *item* dari satu set *item* pendukung (transaksi yang mencakup semua *item*). Setelah menerima satu set *item*, dari satu set *item*, jika dukungan minimum terpenuhi, dukungan minimum akan terlampaui. 1-*Item set* adalah pola yang sangat umum.

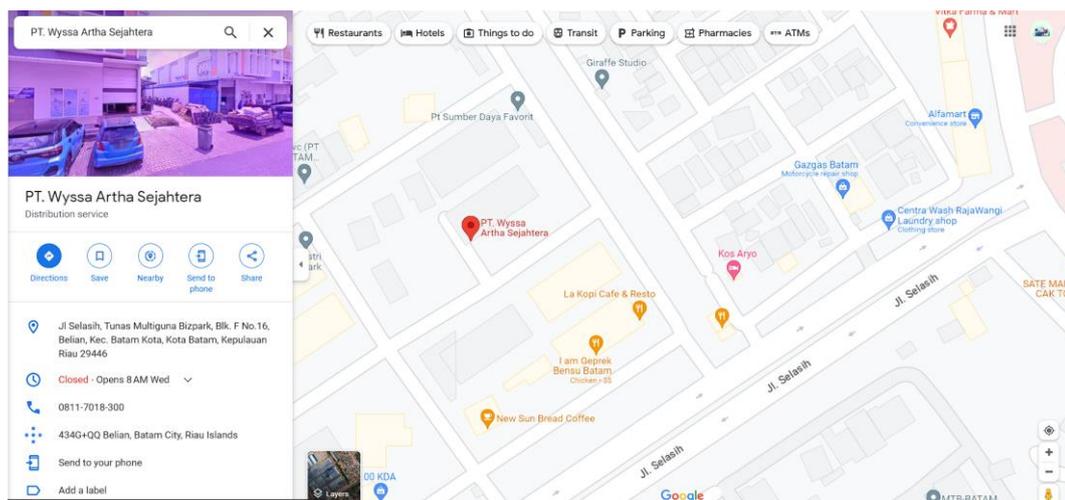
Iterasi 2: Untuk mendapatkan dua set *item*, Anda perlu menjalankan kombinasi set *item* k sebelumnya dan kemudian memindai ulang *database* untuk menghitung *item* yang berisi dukungan. Kumpulan *item* yang mampu mengisi nilai minimum dipilih sebagai pola paling umum untuk kandidat

Iterasi 3 : Atur nilai k-*itemset* dari *support* yang telah mencapai minimum *support* untuk k-*itemset*. Jalankan proses berulang berikut hingga kehabisan k *itemset* yang memenuhi dukungan minimum.

## 3.5 Lokasi dan Jadwal Penelitian

### 3.5.1 Lokasi Penelitian

Lokasi peneliti melakukan penelitian berada di Batam yakni di kantor PT Wyssa Artha Sejahtera Jl Selasih, Tunas Multiguna Bizpark, Blk. F No.16, Belian, Kec. Batam Kota, Kota Batam, Kepulauan Riau 29446 dan Gudangnya yang terletak di Jl. Sudirman Industrial Area, Sukajadi, Kec. Batam Kota, Kota Batam, Kepulauan Riau 29444. Alamat detail yang berdasarkan google map dapat dilihat sesuai dengan letak titik yang menggunakan aplikasi *Google Map* sebagai berikut.



Gambar 3.3 Lokasi Penelitian

### 3.5.2 Jadwal Penelitian

Dalam penelitian yang dilakukan oleh peneliti berdasarkan jadwal yang telah dirancang secara sistematis agar penulisan dapat diselesaikan oleh penulis dalam waktu yang tepat sebagai berikut

| No | Kegiatan              | Waktu Kegiatan   |   |   |   |               |   |   |   |               |   |   |   |             |   |   |   |              |   |   |   |
|----|-----------------------|------------------|---|---|---|---------------|---|---|---|---------------|---|---|---|-------------|---|---|---|--------------|---|---|---|
|    |                       | Februari<br>2023 |   |   |   | Maret<br>2023 |   |   |   | April<br>2023 |   |   |   | Mei<br>2023 |   |   |   | Juni<br>2023 |   |   |   |
|    |                       | 1                | 2 | 3 | 4 | 1             | 2 | 3 | 4 | 1             | 2 | 3 | 4 | 1           | 2 | 3 | 4 | 1            | 2 | 3 | 4 |
| 1  | Pengajuan<br>Judul    |                  |   |   |   |               |   |   |   |               |   |   |   |             |   |   |   |              |   |   |   |
| 2  | Penyusunan<br>BAB I   |                  |   |   |   |               |   |   |   |               |   |   |   |             |   |   |   |              |   |   |   |
| 3  | Penyusunan<br>BAB II  |                  |   |   |   |               |   |   |   |               |   |   |   |             |   |   |   |              |   |   |   |
| 4  | Penyusunan<br>BAB III |                  |   |   |   |               |   |   |   |               |   |   |   |             |   |   |   |              |   |   |   |
| 5  | Penyusunan<br>BAB IV  |                  |   |   |   |               |   |   |   |               |   |   |   |             |   |   |   |              |   |   |   |

Table 3.1 *Jadwal Penelitian*

Sumber: Peneliti 2023