

**PENERAPAN DATA *MINING* UNTUK MENGANALISA DATA  
PENJUALAN BARANG DI SWALAYAN DENGAN MENGGUNAKAN  
ALGORITMA FP-GROWTH**

**SKRIPSI**



**FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS BUDDHI DHARMA  
TANGERANG**

**2022**

**PENERAPAN DATA *MINING* UNTUK MENGANALISA DATA  
PENJUALAN BARANG DI SWALAYAN DENGAN MENGGUNAKAN  
ALGORITMA FP-GROWTH**

**SKRIPSI**

**Diajukan sebagai salah satu syarat untuk kelengkapan gelar kesersarjanaan pada  
Program Studi Teknik Informatika**

**Jenjang Pendidikan Strata 1**



**Disusun oleh:**

**Thomas Febrian**

**20181000037**

**TEKNIK INFORMATIKA**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS BUDDHI DHARMA**

**TANGERANG**

**2022**

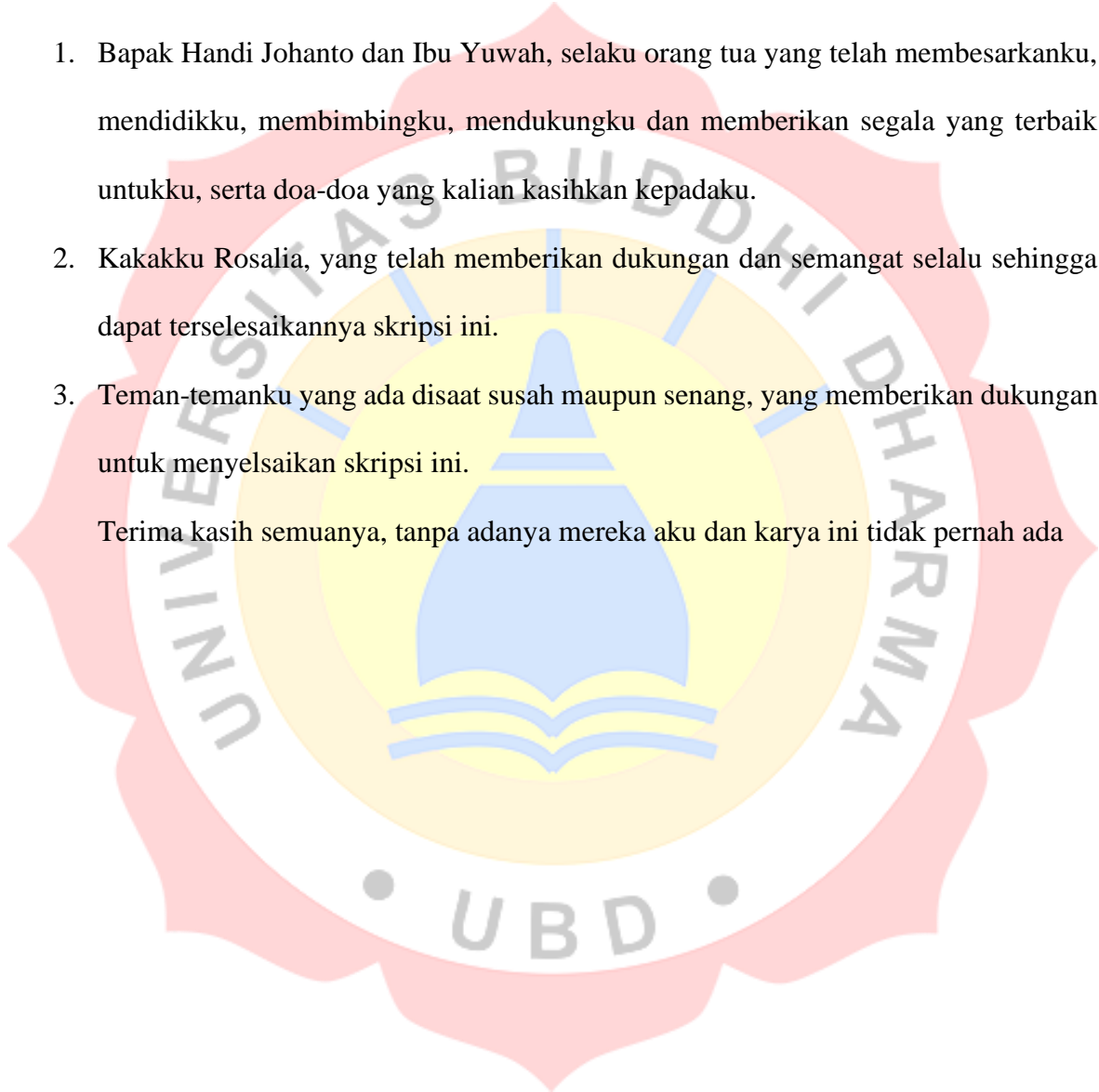
## LEMBAR PERSEMBAHAN

*“The purpose of our lives is to be happy.” – Dalai Lama*

Dengan mengucap puji dan syukur kepada Tuhan Yang Maha Esa, Skripsi ini kupersembahkan untuk:

1. Bapak Handi Johanto dan Ibu Yuwah, selaku orang tua yang telah membesarkanku, mendidikku, membimbingku, mendukungku dan memberikan segala yang terbaik untukku, serta doa-doa yang kalian kasihkan kepadaku.
2. Kakakku Rosalia, yang telah memberikan dukungan dan semangat selalu sehingga dapat terselesaikannya skripsi ini.
3. Teman-temanku yang ada disaat susah maupun senang, yang memberikan dukungan untuk menyelesaikan skripsi ini.

Terima kasih semuanya, tanpa adanya mereka aku dan karya ini tidak pernah ada



**UNIVERSITAS BUDDHI DHARMA**  
**LEMBAR PERNYATAAN KEASLIAN SKRIPSI**

Yang bertanda tangan dibawah ini,

NIM : 20181000037  
Nama : Thomas Febrian  
Jenjang Studi : Strata 1  
Program Studi : Teknik Informatika  
Peminatan : *Database Development*



Dengan ini, saya menyatakan bahwa:

1. Skripsi ini adalah skripsi asli yang belum pernah diajukan untuk mendapatkan gelar akademik, Baik di Universitas Buddhi Dharma maupun di perguruan tinggi atau kampus lainnya.
2. Skripsi ini dibuat sendiri tanpa ada bantuan pihak lain, terkecuali arahan dari dosen pembimbing.
3. Dalam skripsi ini tidak terdapat karya ataupun pendapat yang ditulis maupun dipublikasi orang lain kecuali, tertulis jelas dan juga dicantumkan.
4. Dalam skripsi ini tidak terdapat pemalsuan seperti artikel, jurnal, buku serta tanda tangan yang ada didalamnya.
5. Lembar pengesahan ini saya buat dengan sesungguhnya tanpa ada paksaan.

Sekali lagi saya menyatakan bahwa, skripsi ini saya buat dengan sungguh-sungguh tanpa ada paksaan atau unsur lain didalamnya.

Tangerang, 01 Agustus 2022

Yang menyatakan..



Thomas Febrian

20181000037

## UNIVERSITAS BUDDHI DHARMA

### LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH

NIM : 20181000037  
Nama : Thomas Febrian  
Jenjang Studi : Strata 1  
Program Studi : Teknik Informatika  
Peminatan : *Database Development*



Dengan ini saya menyampaikan bahwa, demi pengembangan ilmu pengetahuan generasi berikutnya, saya Hansen menyetujui serta memberikan izin kepada pihak kampus Universitas Buddhi Dharma atas hak – hak bebas royalti dan sebagainya atas skripsi yang saya buat dengan judul “**PENERAPAN DATA MINING UNTUK MENGANALISA DATA PENJUALAN BARANG DI SWALAYAN DENGAN MENGGUNAKAN ALGORITMA FP-GROWTH**” untuk disimpan, dikelola didalam penyimpanan data serta dipublikasikannya di internet atau media lainnya untuk kepentingan akademik kedepannya selama serta mencantumkan nama saya sebagai pemilik hak cipta.

Sekali lagi saya sampaikan bahwa pernyataan ini saya buat dengan sungguh-sungguh tanpa ada paksaan dari pihak lain ataupun hal lainnya.

Tangerang, 01 Agustus 2022

Yang menyatakan,



Thomas Febrian

20181000037

**UNIVERSITAS BUDDHI DHARMA**  
**LEMBAR PENGESAHAN PEMBIMBING**  
**PENERAPAN DATA MINING UNTUK MENGANALISA DATA**  
**PENJUALAN BARANG DI SWALAYAN DENGAN MENGGUNAKAN**  
**ALGORITMA FP-GROWTH**

Dibuat Oleh:

NIM : 20181000037  
Nama : Thomas Febrian



Telah disetujui untuk dipertahankan di hadapan Tim Penguji Ujian Komprehensif

Program Studi Teknik Informatika

Peminatan *Database Development*

Tahun Akademik 2021/2022

Disahkan oleh,

Tangerang, 22 Juli 2022

Pembimbing,

A handwritten signature in black ink, appearing to be 'Rino, M. Kom', written over a white background.

**Rino, M.Kom**

**NIDN. 0420058502**

**UNIVERSITAS BUDDHI DHARMA**  
**LEMBAR PENGESAHAN SKRIPSI**  
**PENERAPAN DATA MINING UNTUK MENGANALISA DATA**  
**PENJUALAN BARANG DI SWALAYAN DENGAN MENGGUNAKAN**  
**ALGORITMA FP-GROWTH**



Dibuat Oleh:

NIM : 20181000037  
Nama : Thomas Febrian

Telah disetujui untuk dipertahankan di hadapan Tim Penguji Ujian  
Koomprehensif

Program Studi Teknik Informatika  
Peminatan Database Development  
2022  
Disahkan oleh,

Tangerang, 01 Agustus 2022

Dekan,

(Dr. Eng, Ir. Amin Suyitno, M.Eng.)

NIDK: 8826333420

Ketua Program Studi,

(Hartana Wijaya, S.Kom., M.Kom.)

NIDN: 0412058102

**LEMBAR PENGESAHAN TIM PENGUJI**



Nama : Thomas Febrian  
NIM : 20181000037  
Fakultas : Sains dan Teknologi  
Judul : **PENERAPAN DATA MINING UNTUK MENGANALISA DATA  
PENJUALAN BARANG DI SWALAYAN DENGAN  
MENGUNAKAN ALGORITMA FP-GROWTH**

Dinyatakan LULUS setelah mempertahankan didepan Tim Penguji pada hari Senin, 01  
Agustus 2020

Nama Penguji:

Tanda Tangan:

Ketua Sidang : Desiyanna Lasut, M.Kom  
NIDN. 0402128601

Penguji I : Ramona Dyah Safitri, S.Si, M.Si  
NIDN. 0420039301

Penguji II : Rino, M.Kom  
NIDN. 0420058502

Mengetahui,

Dekan Fakultas Sains dan Teknologi

**Dr. Eng, Ir. Amin Suyitno, M.Eng**  
**NIDK: 8826333420**



## KATA PENGANTAR

Dengan mengucapkan Puji Syukur kepada Tuhan Yang Maha Esa, yang telah memberikan Rahmat dan karunia-Nya kepada penulis sehingga dapat menyusun dan menyelesaikan Skripsi ini dengan judul “Penerapan Data Mining Untuk Menganalisa Data Penjualan Barang Di Swalayan Dengan Menggunakan Algoritma FP-Growth” tepat pada waktunya. Yang bertujuan sebagai salah satu syarat untuk menyelesaikan program studi pendidikan Strata 1 (S1) pada program studi Teknik Informatika Universitas Buddhi Dharma. Dalam penulisan ini, penulis banyak mendapatkan bantuan dan dorongan baik moril maupun materiil dari berbagai pihak, maka pada kesempatan kali ini penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Bapak Dr. Suryadi Winata, SE, MM, M.Si., Ak., CA, CMA,CBV, ACPA, CPA (Aust.), sebagai pelaksana tugas rektor Universitas Buddhi Dharma.
2. Bapak Dr. Eng. Ir. Amin Suyitno, M.Eng., sebagai Dekan Fakultas Sains dan Teknologi Universitas Buddhi Dharma
3. Bapak Hartana Wijaya, S.Kom., M.Kom, sebagai Ketua Program Studi Teknik Informatika Universitas Buddhi Dharma
4. Bapak Rino, M.Kom, sebagai Dosen Pembimbing Skripsi yang telah membantu memberikan dukungan dan harapan untuk menyelesaikan skripsi ini.
5. Bapak dan Ibu Dosen Universitas Buddhi Dharma yang telah mengajarkan materi-materi.
6. Keluarga yang selalu memberikan dukungan baik moril maupun materiil serta doa yang diberikan.
7. Teman-teman yang memberikan dukungan dan semangat.

Serta semua pihak yang tidak dapat disebutkan satu persatu sehingga terwujudnya penulisan ini. Penulis menyadari bahwa penyusunan skripsi ini masih jauh dari kata sempurna, maka dari itu penulis mengharapkan adanya kritik dan saran untuk dapat memperbaiki Penulisan ini dimasa yang mendatang.

Akhir kata semoga penulisan skripsi ini dapat berguna bagi penulis dan bagi pembaca serta dapat mendapatkan manfaat dari skripsi ini

Tangerang, 01 Agustus 2022

Penulis

# Penerapan Data Mining Untuk Menganalisa Data Penjualan Barang Di Swalayan Dengan Menggunakan Algoritma FP-Growth

82 + xi halaman/ 10 tabel/ 31 gambar/ 1 lampiran)

## ABSTRAK

Transaksi penjualan yang dilakukan setiap hari akan mengakibatkan terjadinya penumpukan data, karena penjualan yang dilakukan setiap hari akan menyebabkan banyaknya transaksi penjualan. Salah satu strategi untuk dapat meningkatkan penjualan yang dilakukan adalah dengan memanfaatkan data penjualan yang telah ada. Data tersebut digunakan bukan hanya untuk arsip perusahaan saja, tetapi dapat diolah dan dimanfaatkan untuk menghasilkan suatu informasi yang berguna untuk meningkatkan penjualan suatu produk. Pemanfaatan data di dalam suatu sistem informasi untuk mendorong kegiatan pengambilan keputusan tidak cukup hanya mengandalkan data operasional saja, tetapi diperlukan suatu analisis data untuk menggali potensi-potensi informasi dari data yang ada. Para pengambil keputusan berusaha untuk memanfaatkan data-data yang dimiliki dalam mengambil keputusan di dalam gudang data, hal ini mendorong munculnya cabang ilmu baru untuk mengatasi masalah penggalian informasi atau pola penting dari suatu data yang berjumlah besar, yang disebut dengan *data mining*. Salah satu metode yang dapat menggali informasi ini adalah metode asosiasi dengan algoritma FP-Growth. Algoritma ini merupakan salah satu metode dalam penambangan data yang bertujuan untuk mengetahui produk-produk yang sering terjual dan keterkaitan antar produk satu dengan yang lainnya. Penelitian ini memanfaatkan data sekunder berasal dari situs kaggle yang berjudul *groceries dataset* dan data diambil berdasarkan atribut yang diperlukan. Pola penjualan kombinasi barang swalayan ini ditentukan berdasarkan nilai support dan nilai *confidence*, sehingga berdasarkan hasil yang diperoleh nilai support terbesar pada item *whole milk* dan support terkecil pada item *female sanitary product*. Dan gabungan 3 barang dengan *confidence* tertinggi yaitu *domestic egg* dan *meat* kepada *whole milk*.

Kata Kunci: *Data mining*, metode asosiasi, algoritma FP-Growth, *rapidminer*

*Application of Data Mining to Analyze Sales of Goods in Supermarkets Using the FP-Growth Algorithm*

82 + xi pages/ 10 tables/ 31 pictures/ 1 library)

**ABSTRACT**

*Sales transactions made every day will result in data accumulation, because sales made every day will cause many sales transactions. One strategy to increase sales is to utilize existing sales data. The data is used not only for company records, but can be processed and used to produce useful information to increase sales of a product. Utilization of data in an information system to encourage decision-making activities is not enough to rely solely on operational data, but a data analysis is needed to explore the potential of information from existing data. Decision makers try to take advantage of the data they have in making decisions in the data warehouse, this encourages the emergence of new branches of science to overcome the problem of extracting important information or patterns from large amounts of data, which is called data mining. can dig this information is the method of association with the FP-Growth algorithm. This algorithm is one of the methods in data mining that aims to find out the products that are often sold and the relationship between products with one another. This study utilizes secondary data from the kaggle site entitled groceries dataset and the data is taken based on the required attributes. The sales pattern of this combination of self-service goods is determined based on the support value and confidence value, so that based on the results obtained the largest support value is for whole milk items and the smallest support is for female sanitary product items. And a combination of 3 items with the highest confidence, namely domestic egg and meat to whole milk.*

*Keywords: Data mining, association method, FP-Growth algorithm, rapidminer*

## DAFTAR ISI

LEMBAR JUDUL DALAM SKRIPSI

LEMBAR PERNYATAAN KEASLIAN SKRIPSI

LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH

LEMBAR PENGESAHAN PEMBIMBING

LEMBAR PENGESAHAN PENGUJI SKRIPSI

LEMBAR PENGESAHAN SKRIPSI

<b>KATA PENGANTAR</b> .....	viii
<b>ABSTRAK</b> .....	ix
<b>ABSTRACT</b> .....	x
<b>DAFTAR ISI</b> .....	xi
<b>Daftar Tabel</b> .....	xv
<b>Daftar Gambar</b> .....	xvi
<b>Daftar Lampiran</b> .....	xvii
<b>BAB I</b> .....	1
<b>PENDAHULUAN</b> .....	1
1.1 Latar Belakang .....	1
1.2 Identifikasi Masalah .....	2
1.3 Rumusan Masalah .....	3
1.4 Tujuan dan Manfaat Penelitian .....	3
1.4.1 Tujuan Penelitian .....	3
1.4.2 Manfaat Penelitian .....	3
1.5 Ruang Lingkup .....	4
1.6 Metodologi Penelitian .....	4
1.6.1 Metode Penelitian .....	4

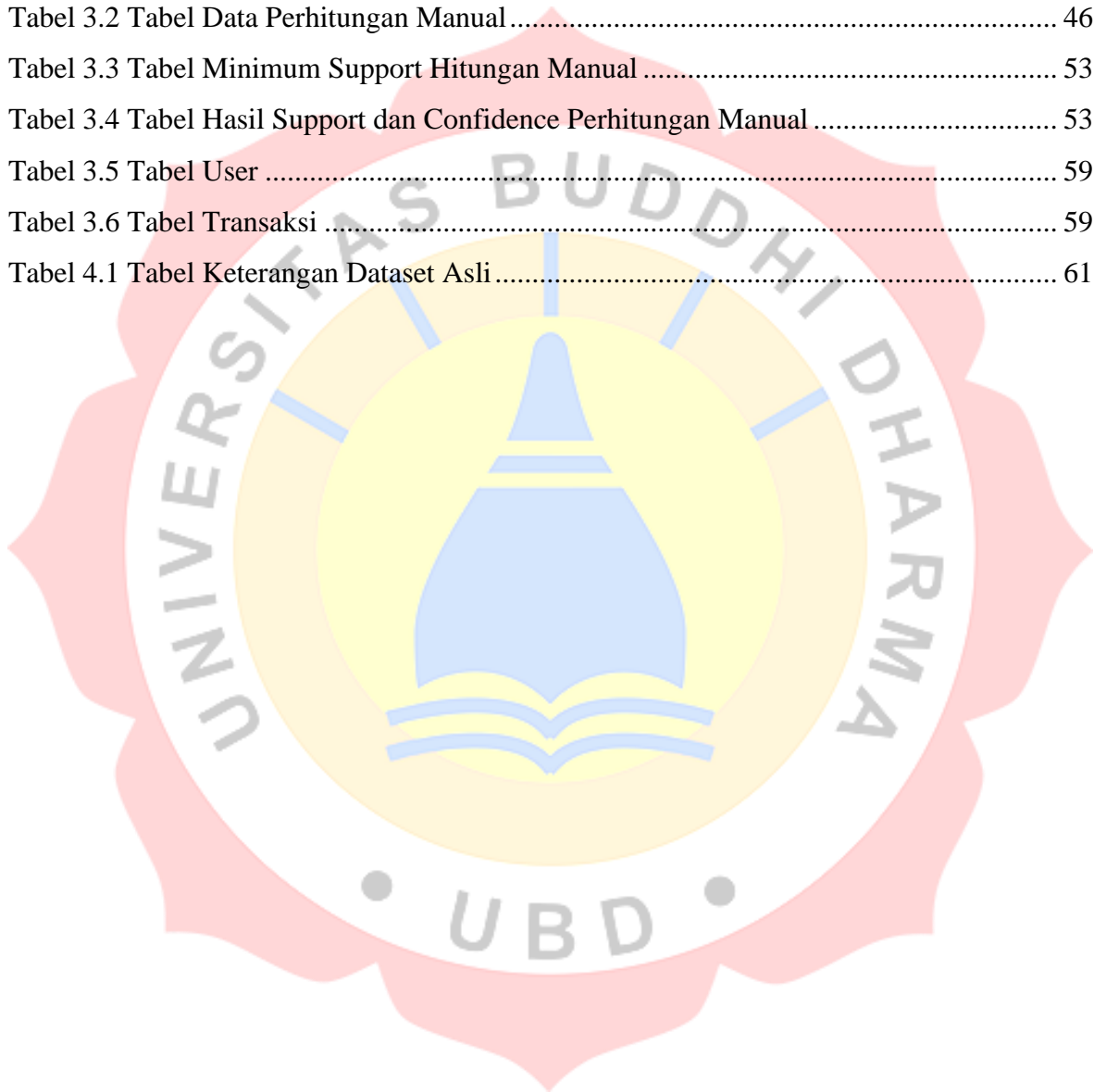
1.6.2 Teknik Pengumpulan Data.....	5
1.7. Sistematika Penulisan.....	5
<b>BAB II.....</b>	<b>7</b>
<b>LANDASAN TEORI.....</b>	<b>7</b>
2.1. Teori Umum.....	7
2.1.1. Data.....	7
2.1.2. Sistem.....	8
2.1.3. Aplikasi.....	8
2.1.4. Informasi.....	8
2.2. Teori Khusus.....	9
2.2.1. Prediksi/ <i>Forecasting</i> .....	9
2.2.2. Penjualan.....	9
2.2.3. Swalayan.....	10
2.2.4. <i>Knowledge Discovery in Database</i> .....	10
2.2.5. Asosiasi.....	11
2.2.6. Algoritma.....	12
2.2.7. FP-Growth.....	13
2.2.8. Data Mining.....	14
2.2.8. CRISP-DM.....	18
2.3. Teori Analisa dan Perancangan.....	19
2.3.1. Rapid Miner.....	19
2.3.2. Netbeans IDE.....	21
2.3.3. Java.....	22
2.3.4. Xampp.....	23
2.3.5. Database.....	25
2.3.6. UML.....	26
2.3.7. ERD.....	27

2.3.8. <i>Black Box Testing</i> .....	28
2.3.9. <i>MySql</i> .....	30
2.4. Tinjauan Studi .....	31
2.4.4 Tabel Perbandingan Jurnal.....	38
2.5. Kerangka Pemikiran .....	41
<b>BAB III</b> .....	43
<b>ANALISA MASALAH</b> .....	43
3.1 Analisa Kebutuhan.....	43
3.2 Konstruksi Algoritma dan Metode .....	45
3.3 Perancangan Layar, Menu, Database.....	53
3.3.1 Perancangan Halaman <i>Login</i> .....	53
3.3.2 Perancangan Halaman Registrasi .....	54
3.3.3 Perancangan Halaman Menu Utama .....	55
3.3.4 Perancangan Halaman Menu Input Data .....	56
3.3.5 Perancangan Halaman Menu Perhitungan.....	57
3.3.6 Perancangan Database .....	58
<b>BAB IV</b> .....	60
<b>PENGUJIAN DAN IMPLEMENTASI</b> .....	60
4.1 Pembahasan Metode dan Algoritma.....	60
4.1.1 <i>Business Understanding</i> .....	60
4.1.2 <i>Data Understanding</i> .....	61
4.1.3 <i>Data Preparation</i> .....	62
4.1.4 <i>Modeling</i> .....	65
4.1.5 <i>Evaluation</i> .....	66
4.1.6 <i>Deployment</i> .....	67
4.2 Spesifikasi <i>Hardware</i> Dan <i>Software</i> .....	70
4.2.1 Spesifikasi <i>Hardware</i> .....	70

4.2.2 Spesifikasi <i>Software</i> .....	70
4.3 Tampilan Program .....	70
4.3.1 Tampilan Halaman Login .....	70
4.3.2 Tampilan Halaman Register .....	71
4.3.3 Tampilan Halaman Menu Utama.....	72
4.3.4 Halaman Menu Input .....	73
4.3.5 Halaman Menu Hasil Perhitungan.....	74
4.6 Pengujian Aplikasi.....	75
4.6.1 Pengujian form login .....	76
4.6.2 Pengujian form registrasi .....	77
4.6.3 Pengujian menu utama.....	77
4.6.4 Pengujian menu input data.....	79
4.6.5 Pengujian menu hitung hasil asosiasi .....	80
<b>BAB V</b> .....	<b>81</b>
<b>KESIMPULAN DAN SARAN</b> .....	<b>81</b>
5.1. Kesimpulan.....	81
5.2. Saran .....	81
DAFTAR PUSTAKA.....	82
DAFTAR RIWAYAT HIDUP .....	84

## Daftar Tabel

Tabel 2.1 Tabel Jurnal Referensi ke 1 .....	31
Tabel 2.2 Tabel Jurnal Referensi ke 2 .....	33
Tabel 2.3 Tabel Referensi ke 3 .....	34
Tabel 3.1 Tabel Keterangan Data .....	44
Tabel 3.2 Tabel Data Perhitungan Manual .....	46
Tabel 3.3 Tabel Minimum Support Hitungan Manual .....	53
Tabel 3.4 Tabel Hasil Support dan Confidence Perhitungan Manual .....	53
Tabel 3.5 Tabel User .....	59
Tabel 3.6 Tabel Transaksi .....	59
Tabel 4.1 Tabel Keterangan Dataset Asli .....	61

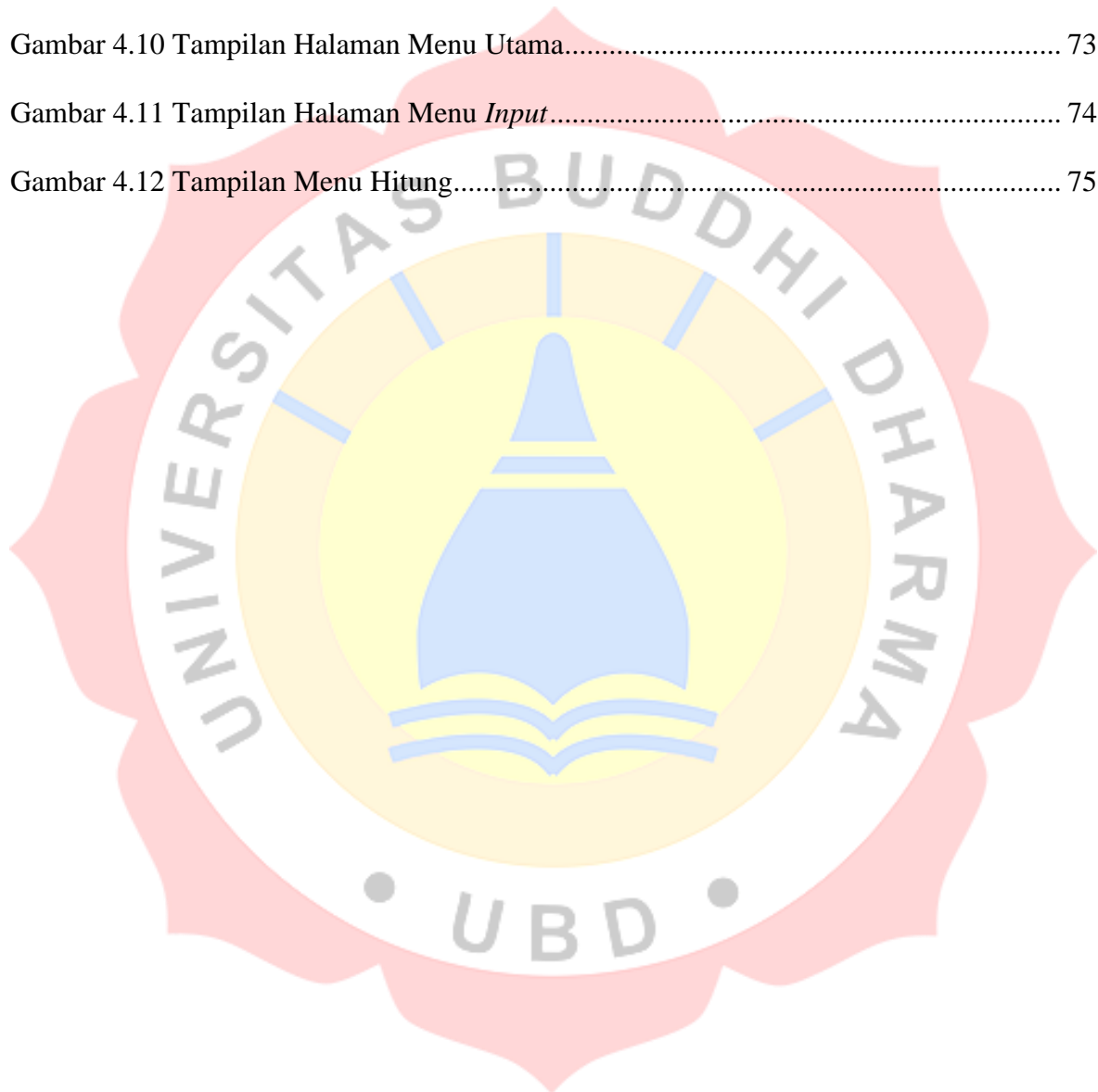




## Daftar Gambar

Gambar 2.1 Kerangka Pemikiran .....	41
Gambar 3.1 Gambar Data Set Asli .....	43
Gambar 3.2 Gambar Data Set Cleaning .....	45
Gambar 3.3 Gambar Tabel Item Support .....	46
Gambar 3.4 Gambar Nilai Prioritas Item.....	47
Gambar 3.5 Gambar Urutan Item Berdasarkan Prioritas .....	47
Gambar 3.6 Gambar FP Tree Langkah Pertama.....	48
Gambar 3.7 Gambar FP Tree Langkah Kedua .....	49
Gambar 3.8 Gambar FP Tree Langkah Ketiga .....	50
Gambar 3.9 Gambar FP Tree Langkah Keempat .....	50
Gambar 3.10 Gambar FP Tree Langkah Kelima.....	51
Gambar 3.11 Gambar FP Tree Langkah Keenam .....	51
Gambar 3.12 Gambar FP Tree Langkah Ketujuh.....	52
Gambar 3.13 Gambar FP Tree Langkah Kedelapan.....	52
Gambar 3.14 Gambar Perancangan Halaman Login .....	54
Gambar 3.15 Gambar Perancangan Halaman Registrasi.....	55
Gambar 3.16 Gambar Perancangan Halaman Menu Utama.....	56
Gambar 3.17 Gambar Perancangan Halaman Menu Input.....	57
Gambar 3.18 Gambar Perancangan Halaman Menu Hitung .....	58
Gambar 4.1 Tampilan Dataset pada Rapid Miner .....	63
Gambar 4.2 Tampilan Statistik Dataset pada Rapid Miner .....	63
Gambar 4.3 Gambar <i>Dataset Cleaning</i> .....	64
Gambar 4.4 Pemodelan FP Growth Pada RapidMiner.....	65

Gambar 4.5 Hasil Gabungan 3 Item .....	66
Gambar 4.6 Gambar Hasil <i>Confidence</i> dan <i>Support</i> .....	67
Gambar 4.7 Hasil <i>Support</i> Terbesar dan Terkecil .....	67
Gambar 4.8 Tampilan Halaman Menu <i>Login</i> .....	71
Gambar 4.9 Tampilan Halaman Menu <i>Register</i> .....	72
Gambar 4.10 Tampilan Halaman Menu Utama .....	73
Gambar 4.11 Tampilan Halaman Menu <i>Input</i> .....	74
Gambar 4.12 Tampilan Menu Hitung .....	75



## DAFTAR LAMPIRAN

Lampiran A-1: Data Set yang digunakan.....

Lampiran B-1: Kartu Bimbingan Skripsi.....



# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Semakin ketatnya persaingan dalam dunia perdagangan, menuntut para pemilik usaha untuk menemukan cara atau strategi yang dapat meningkatkan penjualan. Transaksi penjualan yang dilakukan setiap hari akan mengakibatkan terjadinya penumpukan data, karena penjualan yang dilakukan setiap hari akan menyebabkan banyaknya transaksi penjualan. Salah satu strategi untuk dapat meningkatkan penjualan yang dilakukan adalah dengan memanfaatkan data penjualan yang telah ada. Data tersebut digunakan bukan hanya untuk arsip perusahaan saja, tetapi dapat diolah dan dimanfaatkan untuk menghasilkan suatu informasi yang berguna untuk meningkatkan penjualan suatu produk.

Menurut (Saputri & Sugiyono, 2021) Dalam dunia bisnis, perusahaan harus dapat bersaing dan berpikir bagaimana caranya agar perusahaan bisa terus berkembang dan dapat memperluas skala bisnisnya tersebut. Agar dapat meningkatkan penjualan produknya, para pelaku usaha harus memiliki berbagai strategi yang dilakukan. Salah satunya yaitu dengan memanfaatkan seluruh data transaksi penjualan yang telah terjadi pada perusahaan itu sendiri. Pemanfaatan data di dalam suatu sistem informasi untuk mendorong kegiatan pengambilan keputusan tidak cukup hanya mengandalkan data operasional saja, tetapi diperlukan suatu analisis data untuk menggali potensi-potensi informasi dari data yang ada. Para pengambil keputusan berusaha untuk memanfaatkan data-data yang dimiliki dalam mengambil keputusan di dalam gudang data, hal ini mendorong munculnya cabang ilmu baru untuk mengatasi masalah penggalian informasi atau pola penting dari suatu data yang berjumlah besar, yang disebut dengan

data *mining*. Penggunaan data mining diharapkan dapat memberikan pengetahuan baru yang sebelumnya tersembunyi di dalam gudang data, sehingga menjadi suatu informasi yang berharga dan dapat dimanfaatkan.

Menurut (Sikumbang, 2018) “*Data mining* merupakan proses *iterative* dan interaktif untuk menemukan pola atau *model* baru yang sempurna, bermanfaat dan dapat

dimengerti dalam suatu *database* yang sangat besar (*massive database*)” dalam mengoperasikan *data mining* dibutuhkan yang namanya algoritma. Salah satu metode yang dapat menggali informasi ini adalah metode asosiasi dengan algoritma FP-Growth. Algoritma ini merupakan salah satu metode dalam penambangan data yang bertujuan untuk mengetahui produk-produk yang sering terjual dan keterkaitan antar produk satu dengan yang lainnya. Algoritma FP-Growth termasuk jenis aturan asosiasi pada data mining yang menyatakan asosiasi antara beberapa atribut, yang sering disebut sebagai *market basket analysis*. Dimana sistem ini nantinya dapat bekerja dengan cara menganalisa dan menemukan pola-pola yang berasosiasi antar produk-produk yang dibeli oleh konsumen.

Oleh karena itu melihat permasalahan di atas maka dilakukan penelitian dengan judul: “Penerapan Data Mining Untuk Menganalisa Data Penjualan Barang Di Swalayan Dengan Menggunakan Algoritma FP-Growth”.

## **1.2 Identifikasi Masalah**

Berdasarkan uraian latar belakang masalah dapat disimpulkan identifikasi masalahnya adalah sebagai berikut:

1. Belum adanya pemanfaatan data transaksi untuk meningkatkan penjualan.
2. Kesulitan dalam mengolah informasi data transaksi sehingga membuat kurang efektif dalam pengelolaan data transaksi penjualan.

### 1.3 Rumusan Masalah

Berdasarkan latar belakang diatas maka rumusan masalah pada penelitian ini adalah:

1. Bagaimana mengimplementasikan *data mining* pada dataset dalam transaksi penjualan pada dataset ini?
2. Bagaimana mencari pola asosiasi atau keterkaitan pada produk transaksi penjualan menggunakan algoritma FP-Growth?

### 1.4 Tujuan dan Manfaat Penelitian

#### 1.4.1. Tujuan Penelitian

Tujuan yang ingin dicapai dari penelitian ini yaitu:

1. Penerapan algoritma FP-Growth dalam mengetahui pola kombinasi item sets yang dapat memudahkan pengusaha dalam mengolah data penjualan.
2. Untuk mengetahui jenis produk yang terjual paling banyak.
3. Menerapkan algoritma FP-Growth pada data transaksi penjualan suatu produk.

#### 1.4.2. Manfaat Penelitian

Adapun manfaat penelitian yang dilakukan yaitu:

1. Mengetahui barang apa saja yang sering di beli oleh konsumen sehingga dapat diperbanyak pengadaannya di gudang.
2. Mengetahui keterkaitan antara satu barang dengan barang lainnya sehingga dapat menempatkan barang barang tersebut di rak yang berdekatan.

3. Sebagai referensi untuk penelitian selanjutnya untuk menerapkan kedalam sistem yang lebih luas dan lebih kompleks.

### **1.5. Ruang Lingkup**

Berdasarkan rumusan masalah yang telah ditetapkan, maka ruang lingkup penelitian ini adalah sebagai berikut.

1. Metode data mining yang digunakan adalah metode asosiasi dengan algoritma FP-Growth.
2. Dataset yang penulis analisa ini difokuskan dari dataset yang diambil melalui website kaggle.com.
3. Penelitian ini menggunakan tools RapidMiner sebagai pendukung hasil akhir dari analisis dengan menggunakan algoritma FP-Growth.

### **1.6. Metodologi Penelitian**

#### **1.6.1 Metode Penelitian**

Metode penelitian pada penelitian ini adalah pendekatan deskriptif analitis.

Langkah-langkah yang perlu dilakukan pada pendekatan deskriptif analitis adalah sebagai berikut:

1. Merumuskan tujuan penelitian.
2. Merumuskan unit studi dan menghubungkan hal yang akan dikaji berdasarkan sifat-sifat yang dimiliki dan proses yang menjadi landasan penelitian.
3. Menentukan rancangan dan Teknik mengumpulkan data.
4. Mengumpulkan data.

5. Mengorganisir informasi, data yang terkumpul dan melakukan interpretasi dan generalisasi.

### **1.6.2 Teknik Pengumpulan Data**

Penelitian ini menggunakan teknik studi Pustaka yaitu mengumpulkan informasi dan data dari berbagai sumber seperti buku, jurnal, artikel dan situs-situs yang dapat dipercaya sesuai dengan topik yang digunakan.

Teknik pengumpulan data yang digunakan pada penelitian ini sebagai berikut:

- a. Teknik Pengumpulan Data Sekunder

Teknik pengumpulan data sekunder yang digunakan oleh penulis adalah data yang disediakan oleh penyedia dataset online melalui website <https://www.kaggle.com/>

- b. Studi Pustaka

Teknik pengumpulan data yang kedua adalah studi pustaka. Studi pustaka bertujuan mencari informasi yang mendukung penelitian ini, baik melalui media seperti buku, internet, maupun media informasi lainnya.

### **1.7. Sistematika Penulisan**

Penelitian ini terdiri dari 5 bab, antara lain:

#### **BAB I PENDAHULUAN**

Bab ini terdiri dari latar belakang, identifikasi masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, metode penelitian, dan sistematika penulisan ini sendiri.

#### **BAB II LANDASAN TEORI**



Bab ini terdiri dari teori umum, teori khusus, teori analisa dan perancangan, tinjauan studi, dan kerangka pemikiran.

### BAB III ANALISA MASALAH & PERANCANGAN APLIKASI

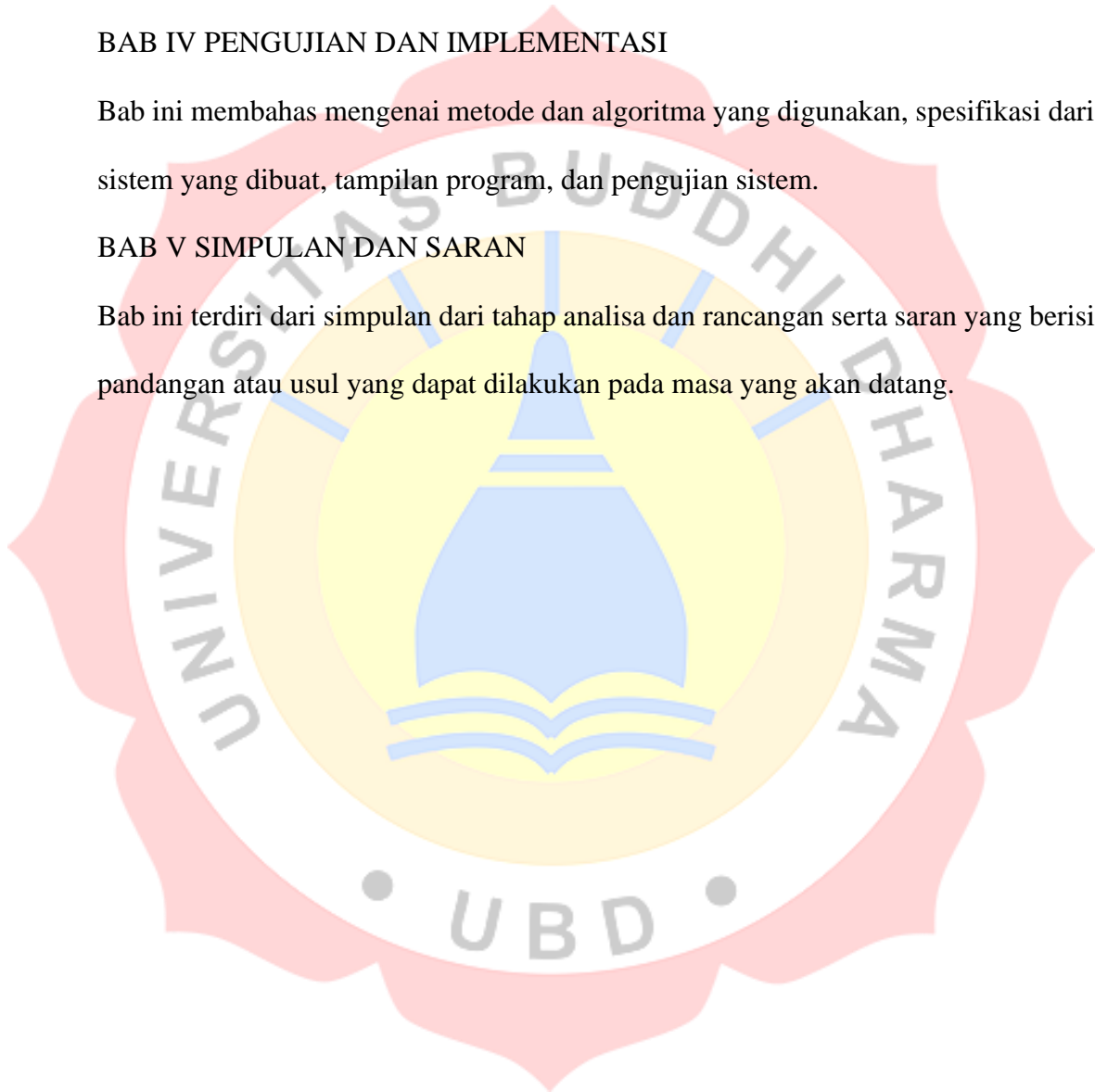
Bab ini memaparkan hasil-hasil dari tahap penelitian, mulai dari analisis, hasil testing, dan implementasinya.

### BAB IV PENGUJIAN DAN IMPLEMENTASI

Bab ini membahas mengenai metode dan algoritma yang digunakan, spesifikasi dari sistem yang dibuat, tampilan program, dan pengujian sistem.

### BAB V SIMPULAN DAN SARAN

Bab ini terdiri dari simpulan dari tahap analisa dan rancangan serta saran yang berisi pandangan atau usul yang dapat dilakukan pada masa yang akan datang.



## **BAB II**

### **LANDASAN TEORI**

#### **2.1. Teori Umum**

##### **2.1.1. Data**

*Data* menurut (Saygılı, 2020) adalah sebuah fakta yang mempresentasikan suatu objek seperti layaknya makhluk hidup (manusia dan hewan) yang dapat dicatat dan disimpan dan memiliki arti yang tersirat.

Menurut (Ladjamudin, 2005), Data adalah komponen dasar dari informasi yang akan diproses lebih lanjut untuk menghasilkan informasi. Sedangkan, menurut Longkutoy dalam buku yang berjudul “Pengenalan komputer”, data adalah suatu istilah majemuk yang berarti fakta atau bagian dari fakta yang mengandung arti yang digabungkan dengan kenyataan, simbol-simbol, gambar-gambar, angka-angka, huruf-huruf, atau simbol-simbol yang menunjukkan suatu ide, objek, kondisi, atau situasi dan lain-lain.

Data sendiri dibagi menjadi 2 yaitu:

a. Kualitatif

Menurut (Jaya, 2020, p. 8) Data Kualitatif adalah data yang didapatkan dari hasil pemberian kode (kategori), dengan isi berupa kata atau didefinisikan sebagai data bukan angka namun diberikan angka.

Contoh: jenis kelamin, status, dan sebagainya. Ciri khas dari data kualitatif adalah tidak dapat melakukan operasi matematika

b. Kuantitatif

Menurut (Jaya, 2020, p. 9) Data Kuantitatif adalah data yang berisikan data nominal atau angka. Data kuantitatif memiliki sifat yang dapat digunakan dalam proses matematika.

**2.1.2. Sistem**

(Subtari, Konsep Dasar Informasi, 2012) berpendapat bahwa sistem merupakan kumpulan dari unsur-unsur dan komponen-komponen yang saling berkaitan dan berhubungan satu sama lain.

Sedangkan menurut (Herliana & Rasyid, 2016) Sistem didefinisikan sebagai suatu kumpulan atau himpunan dari unsur, komponen, atau variabel yang terorganisir, saling interaksi, saling tergantung satu sama lain, dan terpadu.

**2.1.3. Aplikasi**

Menurut (Juansyah, 2015, p. 2) Aplikasi adalah sebuah program utuh yang siap digunakan, yang dibuat untuk membantu seseorang untuk melaksanakan suatu fungsi bagi pengguna jasa aplikasi dan juga pengguna aplikasi lain yang dapat digunakan untuk mencapai suatu tujuan yang ingin dicapai. Aplikasi juga dapat diartikan sebagai sebuah alat bantu yang digunakan oleh orang awam maupun *programmer* dalam menjalankan tugasnya.

**2.1.4. Informasi**

Menurut (Pamungkas, 2017) Informasi merupakan suatu kejadian yang digunakan untuk menggambarkan suatu proses pengambilan keputusan,

yang didapatkan dari hasil pengolahan data, yang dimana informasi mempunyai tingkat yang lebih tinggi, lebih berfungsi serta mempunyai nilai yang lebih berguna dibandingkan data.

Menurut (Supardi, 2015) Informasi adalah data yang telah diklarifikasi

atau diinterpretasi untuk digunakan dalam proses pengambilan keputusan.

Sedangkan menurut (Asmara, 2017, p. 82) “Informasi adalah data yang telah diolah menjadi bentuk yang lebih berguna dan lebih berarti bagi penerimanya.

Data kenyataannya yang menggambarkan suatu kejadian-kejadian”.

## **2.2. Teori Khusus**

### **2.2.1. Prediksi/Forecasting**

Menurut (Nugroho, 2016) Prediksi merupakan kegiatan yang dilakukan oleh seorang peneliti dalam memperkirakan kejadian di masa yang akan datang dengan menggunakan pendekatan ilmu tertentu.

Prediksi/forecasting adalah menentukan jumlah kebutuhan bulan mendatang terkait dengan dukungan data historis (*historical data*) atau serangkaian waktu/periode yang dianalisis sehingga dapat diperhitungkan untuk memprediksi jumlah kebutuhan pada bulan mendatang.

### **2.2.2. Penjualan**

Menurut (Syahrin et al., 2019) Penjualan merupakan sumber hidup suatu perusahaan. Penjualan adalah suatu kegiatan yang dilakukan untuk memindahkan suatu produk dari produsen kepada konsumen demi

mendatangkan keuntungan atau laba dari produk ataupun barang yang dihasilkan produsennya.

### 2.2.3. Swalayan

Supermarket atau pasar swalayan didefinisikan sebagai pranata penjualan eceran yang berstruktur departemental yang menjajakan aneka macam barang dagangan (mencakup makanan dan minuman, daging, sayur mayur, dan produk-produk hasil pabrik pengolahan susu), dan kebanyakan beroperasi atas dasar swalayan dengan pelayanan ke konsumen yang sesedikit mungkin. Jenis barang yang dijual juga semakin berkembang, tidak lagi barang kebutuhan rumah tangga saja tetapi juga obat-obatan, alat-alat kecantikan, buku, majalah, alat-alat musik dan lain-lain. Dari definisi di atas karakteristik supermarket atau pasar swalayan adalah sebagai berikut:

Persediaan barang cukup, Pengaturan barang dibagi kelompok demi kelompok, Harga tertera dengan jelas, Penjaga siap untuk ditanya mengenai jenis dan tempat barang, Parkir untuk membantu kemudahan, Tersedia alat untuk membawa barang-barang yang dibeli, Tempat pembayaran sekaligus siap menyerahkan barang, Untuk keamanan disediakan tempat penitipan tas atau barang tertentu.

### 2.2.4. Knowledge Discovery in Database

*Knowledge Discovery in Database* (KDD) adalah proses menentukan informasi yang berguna serta pola-pola yang ada dalam data. Informasi ini terkandung dalam basis data yang berukuran besar yang sebelumnya tidak diketahui dan potensial bermanfaat. *Data Mining* merupakan salah satu langkah dari serangkaian proses *iterative* KDD (Kusrini, 2007).

Tahapan proses KDD terdiri dari:

1. *Data Selection*. Pada proses ini dilakukan pemilihan himpunan data, menciptakan himpunan data target, atau memfokuskan pada subset *variable* (sampel data) dimana penemuan (*discovery*) akan dilakukan. Hasil seleksi disimpan dalam suatu berkas yang terpisah dari basis data operasional.
2. *Pre-Processing* dan *Cleaning Data*. *Pre-Processing* dan *Cleaning Data* dilakukan membuang data yang tidak konsisten dan *noise*, duplikasi data, memperbaiki kesalahan data dan bisa diperkaya dengan data *eksternal* yang relevan.
3. *Transformation*. Proses ini mentransformasikan atau menggabungkan data ke dalam yang lebih tepat untuk melakukan proses mining dengan cara melakukan peningkatan.
4. *Data Mining*. Proses *data mining* yaitu mencari pola atau informasi menarik dalam data terpilih dengan menggunakan metode atau algoritma tertentu sesuai dengan tujuan dari proses KDD secara keseluruhan.

*Interpretation/Evaluasi*. Proses untuk menerjemahkan pola-pola yang dihasilkan dari *data mining*. Mengevaluasi apakah pola atau informasi yang ditemukan bersesuaian atau bertentangan dengan fakta atau hipotesa sebelumnya. Pengetahuan yang diperoleh dari pola-pola yang terbentuk dipresentasikan dalam bentuk visualisasi.

#### 2.2.5. Asosiasi

Association rules merupakan suatu proses pada data mining untuk menentukan semua aturan asosiatif yang memenuhi syarat minimum untuk support (minsup) dan confidence (minconf) pada sebuah database. Kedua syarat tersebut akan digunakan untuk interesting association rules yang dibandingkan dengan batasan yang telah ditentukan, yaitu minsup dan minconf. (Eska, 2016)

Association rule memiliki dua tahap pengerjaan, yaitu sebagai berikut. (Wijaya, 2018) 1. Mencari frequent itemset 2. Mendefenisikan condition dan result (untuk conditional association rule)

#### 2.2.6. **Algoritma**

Algoritma adalah prosedur yang memiliki tahapan yang logis untuk melakukan pemecahan suatu masalah dengan susunan terstruktur dan sistematis dalam bentuk suatu kalimat dengan jumlah kata-kata yang terbatas. (Marbun et al., 2019).

Algoritma yang baik memiliki beberapa kriteria yang harus diperhatikan sebagai berikut:

a. *Finiteness*

Algoritma yang dikerjakan harus memiliki jeda, algoritma harus berakhir setelah sejumlah langkah yang telah dituliskan.

b. *Definiteness*

Setiap langkah algoritma yang dikerjakan harus didefinisikan dengan tepat, artinya adalah segala langkah yang dilakukan saat menuliskan algoritma haruslah jelas dan tepat, tidak membuat orang lain bingung.

c. *Input*

Setiap algoritma memerlukan data *input* sebagai data yang dapat diolah. Maksudnya adalah setiap algoritma harus memiliki sesuatu data yang perlu dimasukan atau diolah agar nantinya akan menghasilkan suatu hasil atau informasi

d. *Output*

Setiap algoritma setidaknya harus dapat menghasilkan sebuah *output* atau hasil. Suatu algoritma harus memiliki nilai *output* untuk melihat apakah algoritma tersebut berjalan dengan benar atau tidak.

e. *Effectiveness*

Setiap algoritma diharapkan untuk seefisien mungkin, artinya semua operasi dalam algoritma harus cukup mendasar sehingga algoritma tersebut dapat dilakukan secara tepat dan dapat dijalankan dalam jangka waktu yang terbatas.

### 2.2.7. **FP-Growth**

Algoritma FP-Growth merupakan pengembangan dari algoritma Apriori. Algoritma Frequent Pattern Growth (FP-Growth) adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (frequent itemset) dalam sebuah kumpulan data. FP-Growth dapat menemukan frekuensi itemset dengan hanya sedikit mengakses pada database aslinya, dan pendekatannya adalah yang paling efisien. (Wiyana, 2018).



Algoritma FP-Growth termasuk jenis aturan asosiasi pada data mining. Selain FP-Growth, yang termasuk pada golongan ini adalah metode Generalized Rule Induction dan Algoritma Hash Based. Aturan yang menyatakan asosiasi antara beberapa atribut sering disebut affinity analysis atau market basket analysis, (Kusrini & Luthfi, 2009). Langkah-langkah algoritma FP-Growth menurut (Kusrini & Luthfi, 2009) sebagai berikut :

1. Tentukan nilai minimum frequent item set yang menunjukkan itemset memiliki frekuensi kemunculan lebih dari nilai minimum yang telah ditentukan ( $\Phi$ ) untuk mendapatkan kandidat 1 itemset.
2. Setelah mendapatkan hasil dari iterasi ke 1, maka akan dilanjutkan untuk mencari kandidat 2 itemset dan selanjutnya. Proses akan terhenti ketika tidak menemukan yang sesuai dengan nilai minimum yang telah ditentukan.
3. Dari kandidat itemset yang telah diperoleh kemudian dihitung nilai support dan confidence-nya, dengan syarat harus memenuhi nilai minimum support dan confidence yang telah ditentukan.
4. Diperoleh rules yang dapat digunakan sebagai informasi oleh pengguna.

#### **2.2.8. Data Mining**

Menurut (Larose, 2005) Data mining merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang biasanya tidak diduga dan

meringkas data dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemiliknya. Data mining juga merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, dan visualisasi untuk penanganan pengambilan informasi dari database yang besar.

Menurut (Budiarto, 2022) Data Mining adalah langkah analisis terhadap proses penemuan pengetahuan di dalam basis data atau knowledge discovery in databases (KDD). Pengetahuan bisa berupa pola data atau relasi antar data yang valid atau yang tidak diketahui sebelumnya.

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu:

1. **Description (Deskripsi)** Terkadang peneliti dan analis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas pengumpulan suara mungkin tidak dapat menemukan keterangan atau fakta bahwa siapa yang tidak cukup profesional akan sedikit didukung dalam pemilihan presiden. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.
2. **Estimation (Estimasi)** Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numeric daripada ke arah kategori. Model dibangun menggunakan record lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi. Sebagai contoh yaitu estimasi nilai indeks prestasi kumulatif mahasiswa program pasca sarjana dengan melihat nilai indeks prestasi mahasiswa tersebut pada saat mengikuti program sarjana.

3. Prediction (Prediksi) Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada dimasa mendatang. Contoh prediksi dalam bisnis dan penelitian adalah : a. Prediksi harga beras dalam tiga bulan yang akan datang. b. Prediksi tingkat pengangguran lima tahun akan datang. c. Prediski persentase kenaikan kecelakaan lalu lintas tahun depan jika batas bawah kecepatan dinaikan.
4. Classification (Klasifikasi) Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah. Contoh lain klasifikasi dalam bisnis dan penelitian adalah : a. Menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang atau bukan. b. Memperkirakan apakah suatu pengajuan hipotek oleh nasabah merupakan suatu kredit yang baik atau buruk. c. Mendiagnosis penyakit seorang pasien untuk mendapatkan termasuk penyakit apa.
5. Clustering (Pengkusteran) Pengkusteran merupakan pengelompokan record, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan record yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan record-record dalam kluster lain.

Pengklusteran berbeda dengan klasifikasi yaitu tidak adanya variabel target dalam pengklusteran. Pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target. Akan tetapi, algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan record dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan record dalam kelompok lain akan bernilai minimal. Contoh pengklusteran dalam bisnis dan penelitian adalah :

- a. Mendapatkan kelompok-kelompok konsumen untuk target pemasaran dari suatu produk bagi perusahaan yang tidak memiliki dana pemasaran yang besar.
- b. Untuk tujuan audit akuntansi, yaitu melakukan pemisahan terhadap perilaku finansial dalam baik dan mencurigakan.
- c. Melakukan pengklusteran terhadap ekspresi dari gen, untuk mendapatkan kemiripan perilaku dari gen dalam jumlah besar.

6. Association (Asosiasi) Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja. Contoh asosiasi dalam bisnis dan penelitian adalah :

  - a. Meneliti jumlah pelanggan dari perusahaan

telekomunikasi seluler yang diharapkan untuk memberikan respons positif terhadap penawaran upgrade layanan yang diberikan. b. Menemukan barang dalam supermarket yang dibeli secara bersamaan dan barang yang tidak pernah dibeli secara bersamaan.

#### 2.2.8. CRISP-DM

Menurut (Fadillah, 2015) Metodologi CRISP-DM adalah standarisasi data mining yang disusun oleh tiga penggagas data mining market yaitu Daimler Chrysler (Daimler-Benz), SPSS (ISL), NCR. Kemudian dikembangkan pada berbagai workshops (antara 1997-1999). Dalam penerapannya datamining dengan menggunakan metodologi CRISP-DM telah banyak dilakukan, salah satu contohnya dalam jurnal “Business and Data Understanding” dalam rangka pembentukan model tata letak dan tata ruang pasar tradisional dalam rangka meningkatkan laba penjualan menggunakan metode associationrule dan decision tree(studi kasus di kota Bandung)” Jurnal ini membahas mengenai masalah pemahaman tujuan bisnis dan data yang akan digunakan pada proses datamining untuk menentukan tata letak strategi dan tata ruang pasar tradisional dalam rangka meningkatkan laba penjualan.

Menurut (Budiman et al., 2014) Proses data mining berdasarkan CRISP-DM terdiri dari 6 fase, yaitu:

1. *Business Understanding* adalah pemahaman tentang substansi dari kegiatan data *mining* yang akan dilakukan, kebutuhan dari perspektif bisnis.

2. *Data Understanding* adalah fase mengumpulkan data awal, mempelajari data untuk bisa mengenal data yang dipakainya serta mengidentifikasi masalah yang berkaitan dengan data.
3. *Data Preparation* adalah aktivitas yang dilakukan untuk memilih table dan field yang akan ditransformasikan ke dalam *database* yang baru untuk bahan data *mining*.
4. *Modeling* adalah fase menentukan teknik yang dipakai untuk data mining serta menentukan tools yang dipakai untuk melakukan *data mining* dan menentukan algoritma yang dipakai.
5. *Evaluation* adalah fase interpretasi terhadap hasil data mining yang ditunjukkan dalam proses pemodelan pada fase sebelumnya.
6. *Deployment* adalah fase penyusunan laporan atau presentasi dari pengetahuan yang didapat setelah melakukan kegiatan data mining ini.

## 2.3. Teori Analisa dan Perancangan

### 2.3.1. Rapid Miner

Menurut (Jollyta, Siddik, Mawengkang, & Efendi, 2021) *Rapidminer* adalah alat penambangan data yang kuat untuk membuat model prediktif dengan cepat. Alat ini juga menerapkan *all in one* yang dimana menampilkan

raturan persiapan data dan algoritma pembelajaran mesin untuk mendukung semua proyek data mining. *Rapidminer* merupakan paket perangkat lunak yang memungkinkan penambangan data, teks, dan analitik prediktif (*datamining*). Program ini memungkinkan pengguna untuk memasukan data mentah, termasuk *database* dan teks, yang kemudian dianalisis secara otomatis dan cerdas dalam skala besar.

Rapid miner adalah sebuah software data mining untuk melakukan analisis kumpulan data dengan berbagai metode yang dapat digunakan. Rapid miner bersifat open source yang berarti bebas untuk digunakan tanpa harus membayar lisensi rapid miner. RapidMiner dapat memberikan hasil dari berbagai metode deskriptif dan prediksi untuk memberikan informasi bagi pengguna sehingga berisi informasi yang dapat menentukan keputusan terbaik.

Menurut (Munanda & Monalisa, 2021) RapidMiner adalah aplikasi data mining berbasis open-source yang terkemuka dan ternama. Didalamnya terdapat aplikasi yang berdiri sendiri untuk analisis data dan sebagai mesin data mining seperti untuk loading data, transformasi data, pemodelan data, dan metode visualisasi data. RapidMiner pertama kali dinamai Yet Another Learning Environment atau singkat YALE.

Rapidminer mempunyai beberapa fitur yaitu :

1. Banyaknya algoritma data mining, seperti decision tree dan self-organization map.
2. Bentuk grafis yang canggih, seperti tumpang tindih diagram histogram, tree chart dan 3D Scatter plots.

3. Banyaknya variasi plugin, seperti text plugin untuk melakukan analisis teks.

4. Menyediakan prosedur data mining dan machine learning termasuk: ETL (extraction, transformation loading), data preprocessing, visualisasi, modelling dan evaluasi

5. Proses data mining tersusun atas operator-operator yang nestable, dideskripsikan dengan XML, dan dibuat dengan GUI

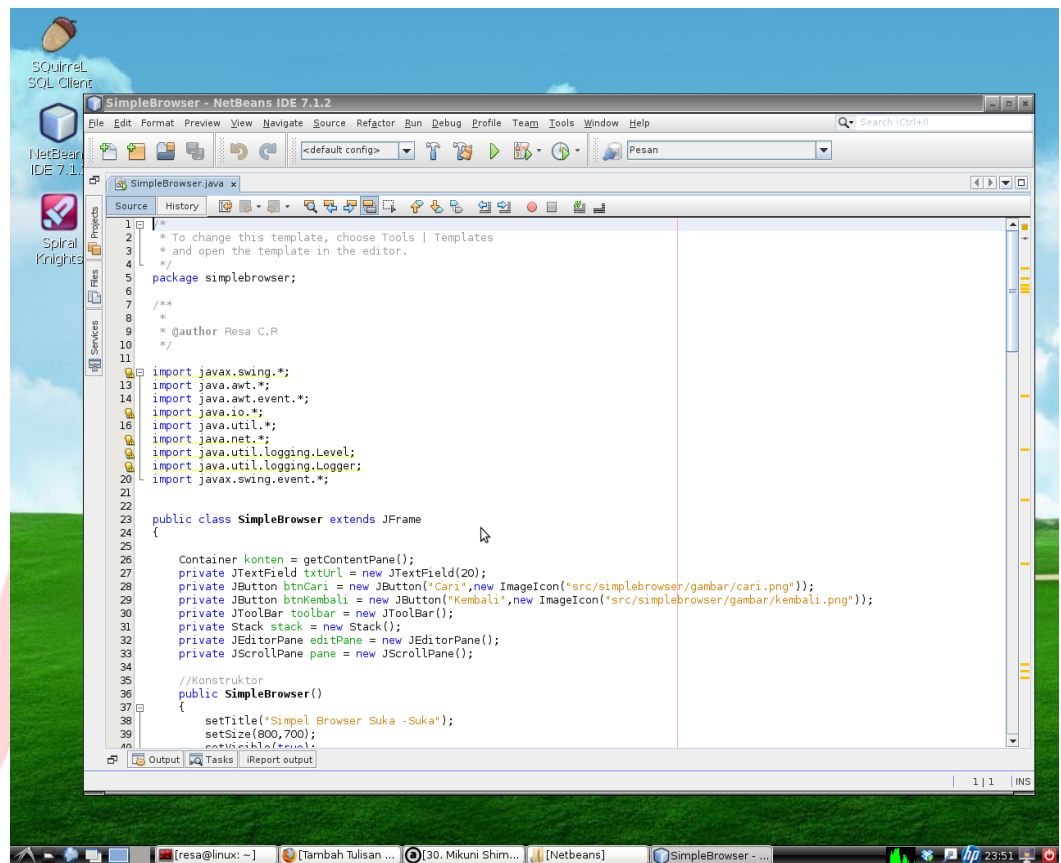
6. Mengintegrasikan proyek data mining Weka.

Rapid Miner juga menerapkan *all in one* yang dimana menampilkan ratusan persiapan data dan algoritma pembelajaran mesin untuk mendukung semua proyek data mining. *Rapidminer* merupakan paket perangkat lunak yang memungkinkan penambahan data, teks, dan analitik prediktif (*datamining*). Program ini memungkinkan pengguna untuk memasukan data mentah, termasuk *database* dan teks, yang kemudian dianalisis secara otomatis dan cerdas dalam skala besar.

### 2.3.2. Netbeans IDE

Menurut (Mallu, 2015) IDE adalah sebuah peralatan dalam pemrograman yang diintegrasikan ke dalam software yang mendukung pembangunan GUI, teks, compiler dan debugger. Perancangan dan pemrograman dilakukan di dalam kerangka sebuah proyek. Proyek NetBeans merupakan sekumpulan file yang dikelompokkan di dalam satu kesatuan. NetBeans IDE dapat digunakan untuk membangun perangkat lunak berbasis *Java Standart Edition*, *Java Enterprise Edition*, *JavaFX*, *PHP*, *C/C++*, *Ruby* dan *Phyton*.





Sumber: <https://computerscience.id/2013/03/mengganti-tampilan-netbeans-ide.html>

Netbeans adalah suatu aplikasi IDE atau Integrated Development Environment yang berbasis bahasa Java dan berjalan diatas Swing. Swing adalah suatu teknologi yang memungkinkan pengembangan aplikasi desktop dan dapat berjalan di berbagai macam platform seperti Windows, Mac OS, Linux dan Solaris

### 2.3.3. Java

Menurut (Jando & Andrianus, 2018) Java adalah bahasa pemrograman yang bersifat OOP (*Object Oriented Programming*) dan dapat dijalankan diberbagai *platform* mana saja seperti, Windows, Linux, DOS dan Unix.

Sedangkan menurut (Andianto, 2017) Java adalah bahasa pemrograman yang dapat dijalankan diberbagai komputer termasuk telepon genggam. Bahasa ini awalnya dibuat oleh James Gosling saat masih bergabung di Sun Microsystems saat ini merupakan bagian dari Oracle dan dirilis tahun 1995. Bahasa ini banyak mengadopsi sintaksis yang terdapat pada C dan C++ namun dengan sintaksis model objek yang lebih sederhana. Java dikenal dengan slogannya “Write once, run anywhere”.

#### 2.3.4. Xampp

Menurut (Fathoroni et al., 2020, p. 50) XAMPP adalah perangkat lunak bebas, yang mendukung banyak sistem operasi, XAMPP merupakan kompilasi dari beberapa program. XAMPP juga merupakan *tools* yang menyediakan paket perangkat lunak kedalam satu buah paket yang terdiri dari Apache, MySql, PHP, dan Pearl, X diartikan sebagai XAMPP dapat diinstal di sistem operasi manapun. XAMPP merupakan suatu *web server* yang mudah digunakan dan dapat melayani tampilan halaman web yang dinamis.



Gambar 2.1 Logo XAMPP

Lima komponen penting dalam XAMPP adalah sebagai berikut.

1. Control panel 31 Layanan yang memberikan akses lebih leluasa dalam mengatur database, mengunggah file, mengelola fitur juga melakukan konfigurasi terkait proyek website.
2. Htdocs Merupakan nama sebuah folder bagian dari XAMPP yang berfungsi untuk menyimpan berbagai file dan dokumen yang akan ditampilkan di website. Kapasitas penyimpanan dalam Htdocs bergantung pada storage harddisk yang ada di komputer.
3. PhpMyAdmin Program untuk mengolah database website yang dapat mengakses PhpMyAdmin dengan berkunjung ke situs localhost/phpMyAdmin.
4. Config Menu yang digunakan untuk melakukan beberapa konfigurasi dasar XAMPP.

5. Netstat Menu yang digunakan untuk memastikan apakah ada aplikasi lain yang menggunakan port XAMPP pada perangkat yang digunakan.

### 2.3.5. Database

Menurut (Indrajani, 2015) basis data adalah kumpulan data yang saling berhubungan secara logis dan didesain untuk mendapatkan data yang dibutuhkan oleh suatu organisasi.

Basis data (*database*) adalah sistem terkomputerisasi yang tujuan utamanya adalah memelihara data yang sudah diolah atau informasi dan membuat informasi tersedia saat dibutuhkan. Pada intinya basis data adalah media untuk menyimpan data agar dapat diakses dengan mudah dan cepat. (Sukanto & Shalahuddin, 2013).

Database mempunyai beberapa jenis yaitu:

1. Analytical database, yang merupakan database untuk menyimpan informasi dan data yang diambil dari operasional dan eksternal database. Database ini terdiri dari data dan informasi yang diringkas dan paling banyak dibutuhkan oleh suatu organisasi manajemen maupun End-user lainnya.

2. Operational database ialah database yang menyimpan data secara rinci yang dibutuhkan untuk mendukung operasi dari seluruh organisasi. Biasa juga disebut dengan SADB (subject-area databases), transaksi, dan produksi database. Contohnya seperti : database inventaris, database pribadi, database pelanggan, akuntansi database.

3. Distributed database merupakan kelompok kerja lokal database dan departemen di kantor-kantor dan lokasi kerja yang lainnya. Dalam

database ini terdapat dua segmen yaitu user database dan operasional yang datanya digunakan dan dihasilkan hanya pada pengguna situs itu sendiri.

4. Data warehouse yaitu sebuah data warehouse yang menyimpan data dari tahun-tahun sebelumnya hingga saat ini. Data warehouse merupakan sumber utama data yang telah terintegrasi sehingga bisa digunakan dan dimanfaatkan oleh para pengguna seluruh organisasi yang profesional. Perkembangan yang terjadi akhir ini dari data warehouse ialah dipergunakan sebagai Shared nothing architecture untuk mendukung dan memfasilitasi ekstrem scaling.

5. End-user database yang terdiri dari file-file data yang dikembangkan dari end-user dalam workstation mereka. Contohnya berbagai koleksi dokumen dalam word processing, spreadsheet maupun download file.

6. Real time database merupakan sebuah sistem pengolahan yang dirancang dalam menangani beban kerja suatu negara yang bisa berubah-ubah, mengandung data terus menerus dan sebagian tidak berpengaruh terhadap waktu. database ini bermanfaat bagi orang-orang hukum, akuntansi, perbankan, multi media dan analisis data yang ilmiah.

#### **2.3.6. UML**

Menurut (Deimas et al., 2021) UML merupakan salah satu tools atau model yang digunakan untuk membuat suatu aplikasi pengembangan software dengan berbasis object-oriented. Selain itu juga mampu menghasilkan tulisan standar pada sistem blue print seperti kerangka bisnis proses, penulisan kelas yang spesifik, skema database, serta

berbagai komponen yang dibutuhkan pada sistem tersebut. Unified Modeling Language merupakan set standar pada teknik untuk membuat diagram. Diagram UML yang paling berguna diantaranya adalah Use Case Diagram, Statechart Diagram, Sequence Diagram, Component Diagram, Class Diagram, Activity Diagram, dan Deployment Diagram.

UML (Unified Modelling Language) adalah suatu metode dalam pemodelan secara visual yang digunakan sebagai sarana perancangan sistem berorientasi objek. Awal mulanya, UML diciptakan oleh Object Management Group dengan versi awal 1.0 pada bulan Januari 1997.

UML juga dapat didefinisikan sebagai suatu bahasa standar visualisasi, perancangan, dan pendokumentasian sistem, atau dikenal juga sebagai bahasa standar penulisan blueprint sebuah software.

UML diharapkan mampu mempermudah pengembangan piranti lunak (RPL) serta memenuhi semua kebutuhan pengguna dengan efektif, lengkap, dan tepat. Hal itu termasuk faktor-faktor scalability, robustness, security, dan sebagainya.

### **2.3.7. ERD**

(Sri Endang Anjarwani & Baiq Maisum Jum, 2022) Mengatakan, Entity Relationship Diagram (ERD) adalah alat pemodelan data utama dan akan membantu mengorganisasi data dalam suatu proyek ke dalam entitas-entitas dan menentukan hubungan antar entitas. ERD Merupakan model relasional yang menyatakan hubungan antar objek data atau entitas.

Langkah-langkah untuk membuat sebuah entity relationship diagram:

1. Lakukan identifikasi dan tentukan seluruh entitas yang akan digunakan

Langkah pertama, yaitu kamu harus melakukan identifikasi untuk menentukan entitas apa saja yang akan kamu gunakan di dalam diagram. Kamu dapat memulai langkah ini dengan membuat sebuah persegi panjang, lalu tuliskan deskripsi singkat mengenai entitas tersebut.

2. Identifikasi dan menjelaskan relasi dari entitas

Langkah berikutnya adalah kamu dapat mencari entitas yang sekiranya memiliki relasi. Kemudian kamu dapat membuat garis yang menghubungkan kedua entitas tersebut, selain itu kamu juga harus menentukan relasi antara satu entitas dengan entitas yang lainnya. Kamu dapat menggunakan simbol belah ketupat untuk mendeskripsikan hubungan relasinya. Di tahap ini juga kamu dapat menjelaskan jenis relasi apa yang digunakan oleh suatu entitas, apakah one to one, one to many, atau many to many.

3. Tambahkan atribut yang diperlukan

Setelah kamu selesai mengidentifikasi dan menjelaskan relasi dari entitas, berikutnya kamu dapat menambahkan atribut-atribut yang sudah kamu pelajari pada artikel ini. Pastikan juga kamu menentukan atribut key pada setiap entitas. Lambangkan atribut key tersebut dengan bentuk oval dan berikanlah deskripsi pada lambang tersebut, ingat ya untuk atribut key deskripsi ditulis dengan garis bawah.

#### 4. Lengkapi diagram

Langkah terakhir yang perlu kamu lakukan untuk menyelesaikan ERD ini adalah melengkapi diagram yang telah kamu buat sesuai dengan kebutuhan dari sistem atau database yang kamu buat. Pada tahap ini kamu harus lebih teliti untuk memeriksa setiap komponen, simbol, dan deskripsi yang salah atau tertukar.

#### 2.3.8. Black Box Testing

Pratama dan Junianto dalam (Rahadi & Vikasari, 2020, hal. 58) berpendapat, “Black box testing cenderung lebih cepat dalam menemukan beberapa hal seperti fungsional sistem yang tidak berjalan dengan semestinya, kesalahan pada struktur data, kesalahan dalam akses database, kesalahan pada tampilan interface, hingga kesalahan pada performa sistem atau aplikasi yang dikembangkan”.

Black Box Testing adalah pengujian yang dilakukan untuk mengamati hasil input dan output dari perangkat lunak tanpa mengetahui struktur kode dari perangkat lunak. Pengujian ini dilakukan di akhir pembuatan perangkat lunak untuk mengetahui apakah perangkat lunak dapat berfungsi dengan baik.

Ada beberapa teknik Black Box Testing diantaranya:

1. All Pair Testing, teknik all pair testing dikenal juga dengan pairwise testing. Pengujian ini digunakan untuk menguji semua kemungkinan kombinasi dari seluruh pasangan berdasarkan input parameteranya.
2. Boundary value analysis



Teknik ini berfokus pada pencarian error dari luar atau sisi dalam perangkat lunak.

### 3. Cause-effect graph

Berikutnya adalah teknik cause-effect graph. Teknik pengujian ini menggunakan grafik sebagai patokannya. Grafik ini menggambarkan relasi antara efek dan penyebab dari error.

### 4. Equivalence partitioning

Teknik ini bekerja dengan cara membagi data input dari beberapa perangkat lunak menjadi beberapa partisi data.

### 5. Fuzzing

Fuzzing merupakan teknik pencarian bug dalam perangkat lunak dengan memasukkan data yang tidak sempurna.

### 6. Orthogonal array testing

Selanjutnya adalah orthogonal array testing. Teknik ini digunakan jika input berukuran kecil, akan tetapi cukup berat jika digunakan dalam skala yang besar.

### 7. State transition

Terakhir adalah state transition. Teknik ini berguna untuk melakukan pengujian terhadap mesin dan navigasi dari UI dalam bentuk grafik.

## 2.3.9. MySql

Menurut (MCLaughlin, 2012, p. 7) MySQL adalah sebuah *database* yang dimana dia menyimpan informasi anda, informasi pengguna, dan informasi lain yang ingin anda masukan. MySQL pada

intinya adalah sebuah gudang yang dimana anda dapat menyimpan barang-barang untuk dilihat nanti, MySQL juga memberikan anda mekanisme yang cepat untuk menemukan semua barang/informasi yang anda simpan digudang kapan pun dibutuhkan. *MySQL* mempunyai fungsi utama untuk mengoperasikan *database*, terutama dalam hal pemilihan, penyelesaian maupun *input data*.

*MySQL* memiliki beberapa perintah dasar dalam pengoperasiannya yaitu:

a. *Create Database*

Perintah ini berfungsi untuk membuat sebuah *database* baru.

b. *Create Table*

Perintah ini berfungsi untuk membuat sebuah *table* baru didalam *database*.

c. *Insert*

Perintah ini berfungsi untuk memasukan data kedalam *table* yang diinginkan.

d. *Delete*

Perintah ini berfungsi untuk menghapus data yang diinginkan pada *table* tertentu

e. *Select*

Perintah ini berfungsi untuk memilih dan membaca data dari *table/database* yang diinginkan.

f. *Update*

Perintah ini berfungsi untuk merubah data pada *table* yang diinginkan.

Sedangkan menurut (Sri Endang Anjarwani & Baiq Maisum Jum, 2022) *MySQL* merupakan aplikasi yang digunakan untuk mengelola

data yang akan tersimpan dalam *database server* bersifat *multiuser* dan *multi-threaded*. *Structured Query Language (SQL)* digunakan untuk mengelola data secara terstruktur dan pada MySQL digunakan untuk menambah, merubah dan menghapus data yang berada dalam database.

## 2.4. Tinjauan Studi

Tabel 2.1 Tabel Jurnal Referensi ke 1

No	Data Jurnal/Makalah	Keterangan
1	Judul	PENERAPAN DATA MINING UNTUK ANALISIS POLA PEMBELIAN KONSUMEN DENGAN ALGORITMA FPGROWTH PADA DATA TRANSAKSI PENJUALAN SPARE PART MOTOR
2	Jurnal	Kumpulan Jurnal Ilmu Komputer (KLIK)
3	Volume dan Halaman	Vol. 05, No. 01
4	Tanggal dan Tahun	Februari 2018
5	Penulis	Alfannisa Annurullah Fajrin, Algifanri Maulana
6	Penerbit	Universitas Putra Batam
7	Tujuan Penelitian	Mengetahui pola pembelian konsumen pada masing-masing cabang, menginterpretasikan pola yang telah dihasilkan menjadi sebuah informasi, merancang bagaimana strategi dan sistem penjualan yang baik, menerapkan metode Data Mining dengan algoritma FP-Growth ke dalam aplikasi analisis pola pembelian konsumen, dan menguji pola yang telah diterapkan berhasil atau tidak.
8	Lokasi dan Subjek Penelitian	CV. TAJAJA BARU

9	Perancangan Sistem	<p>1. Software Microsoft Excel</p> <p>2. Software Rapid Miner</p>
10	Hasil Penelitian	<p>Hasil penjualan dari spare part motor sport yang paling banyak terjual di TBDamar CV.TJAHAJA BARU Padang bisa diketahui dengan menggunakan algoritma FP-Growth. Spare part yang memenuhi minimum support dan minimum confidence serta yang banyak terjual adalah screw valve adjusting, oil seal, battery assy, axle, gasket cylinder, dan cable clutch.</p>
11	Kekuatan Penelitian	<p>Menerapkan metode Data Mining dengan algoritma FP-Growth kedalam aplikasi untuk analisis pola pembelian konsumen sangat bermanfaat bagi perusahaan tersebut, karena TB-Damar akan mengetahui spare part mana yang banyak dibeli dan membantu dalam pemesanan spare part pada kantor pusat.</p>
12	Kelemahan Penelitian	<p>Belum dibuatnya suatu aplikasi perhitungan untuk permasalahan ini</p>
13	Kesimpulan	<p>Menerapkan metode Data Mining dengan algoritma FP-Growth kedalam aplikasi untuk analisis pola pembelian konsumen sangat bermanfaat bagi perusahaan tersebut, karena TB-Damar akan mengetahui spare part mana yang banyak dibeli dan membantu dalam pemesanan spare part pada kantor pusat.</p>

Tabel 2.2 Tabel Jurnal Referensi ke 2

No	Data Jurnal/Makalah	Keterangan
----	---------------------	------------

1	Judul	Penerapan Data Mining dalam Analisis Pola Pembelian Minuman dan Makanan di Kantin SMAN 4 Langsa Menggunakan Algoritma FP-Growth
2	Jurnal	Jurnal Nasional Informatika dan Teknik Jaringan
3	Volume dan Halaman	Vol. 4 No. 2
4	Tanggal dan Tahun	2020
5	Penulis	Rizky Fitria Haya, Desy Ramadani
6	Penerbit	Universitas Samudra
7	Tujuan Penelitian	Mengetahui pola pembelian minuman dan makanan di kantin SMAN 4 Langsa serta merancang bagaimana skema dan sistem penjualan yang baik sehingga dapat menerapkan algoritmaFPGrowth ke dalam aplikasi analisis pola pembelian minuman dan makanan di kantin.
8	Lokasi dan Subjek Penelitian	Kantin SMAN 4 Langsa
9	Perancangan Sistem	Menggunakan aplikasi Rapid Miner dan Microsoft Excel
10	Hasil Penelitian	Dari penelitian yang telah dilakukan pada objek Kantin SMAN 4 Langsa, penjualan produk minuman dan makanan selama 15 hari didapatkan nilai confidence tertinggi adalah 1 atau sekitar 100% dengan aturan (rules) pada setiap pembelian produk Bakso Goreng dapat dipastikan akan membeli Teajus Gula Batu

11	Kekuatan Penelitian	Metode yang digunakan dalam menentukan frequent itemset dari sebuah pohon keputusan yang menggunakan metode algoritma FP-Growth bekerja dengan baik saat melakukan proses pencarian frequent itemset dan proses pembentukan FP-Tree yang menghasilkan rules dari data di kantin SMAN 4 Langsa
12	Kelemahan Penelitian	Tidak disebutkannya kekurangan dalam penelitian ini.
13	Kesimpulan	Penelitian ini berhasil mendapatkan tujuannya dengan menemukan barang mana saja yang keluar secara bersamaan dan kemudian meletakkan barang tersebut dipintu keluar agar lebih efisiensi dalam pengambilan barang dan pencarian barang.

Tabel 2.3 Tabel Referensi ke 3

No	Data Jurnal/Makalah	Keterangan
1	Judul	PENERAPAN ALGORITMA FP-GROWTH PADA DATA TRANSAKSI PENJUALAN UNTUK PENENTUAN TATALETAK BARANG
2	Jurnal	Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi
3	Volume dan Halaman	Vol 7, No. 2 Halaman 173-184
4	Tanggal dan Tahun	Agustus, 2021
5	Penulis	Elvira Munanda, Siti Monalisa
6	Penerbit	Universitas Islam Negeri Sultan Syarif Kasim Riau

7	Tujuan Penelitian	<p>Pada penelitian ini penulis akan mencari item yang sering dibeli bersamaan oleh konsumen dengan menggunakan data transaksi penjualan pada Toko Doa Bunda. Dengan mengetahui pola pembelian item-item yang sering dibeli bersamaan memudahkan pihak pemilik toko dalam menentukan strategi penempatan tata letak barang sesuai dengan kebiasaan konsumen dalam melakukan transaksi</p>
8	Lokasi dan Subjek Penelitian	Toko Doa Bunda, Pekanbaru
9	Perancangan Sistem	Menggunakan aplikasi Rapid Miner dan Microsoft Excel
10	Hasil Penelitian	<p>Setelah dilakukannya penelitian dengan melakukan proses asosiasi untuk mencari rule dengan menggunakan algoritma FPGrowth dapat diketahui item yang sering dibeli secara bersamaan dari bulan April tahun 2017 hingga bulan Agustus tahun 2019.</p> <p>Dari 6366 data transaksi dan data bulan April tahun 2017 sampai Agustus tahun 2019 dapat disimpulkan bahwa item Baju Kaos Pria dan Item Celana pendek pria sering di beli bersamaan, baik menggunakan semua data maupun percobaan perbulan dengan tahun yang berbeda.</p>
11	Kekuatan Penelitian	<p>Dengan adanya hasil analisa ini dapat memberikan informasi kepada pemilik Toko / owner tentang kecendrungan konsumen membeli barang berdasarkan kombinasi 2 itemset sehingga dapat dilakukan pengaturan tata letak barang secara berdekatan untuk mempercepat proses pelayanan dan pencarian item</p>

12	Kelemahan Penelitian	Tidak dibuatnya sistem atau aplikasi dengan menggunakan algoritma FP-Growth.
13	Kesimpulan	Metode yang digunakan terbukti cocok untuk dalam mengasosiasi barang-barang yang dijual, terbukti dengan menerapkan algoritma FP-Growth mendapatkan hasil yang diinginkan.



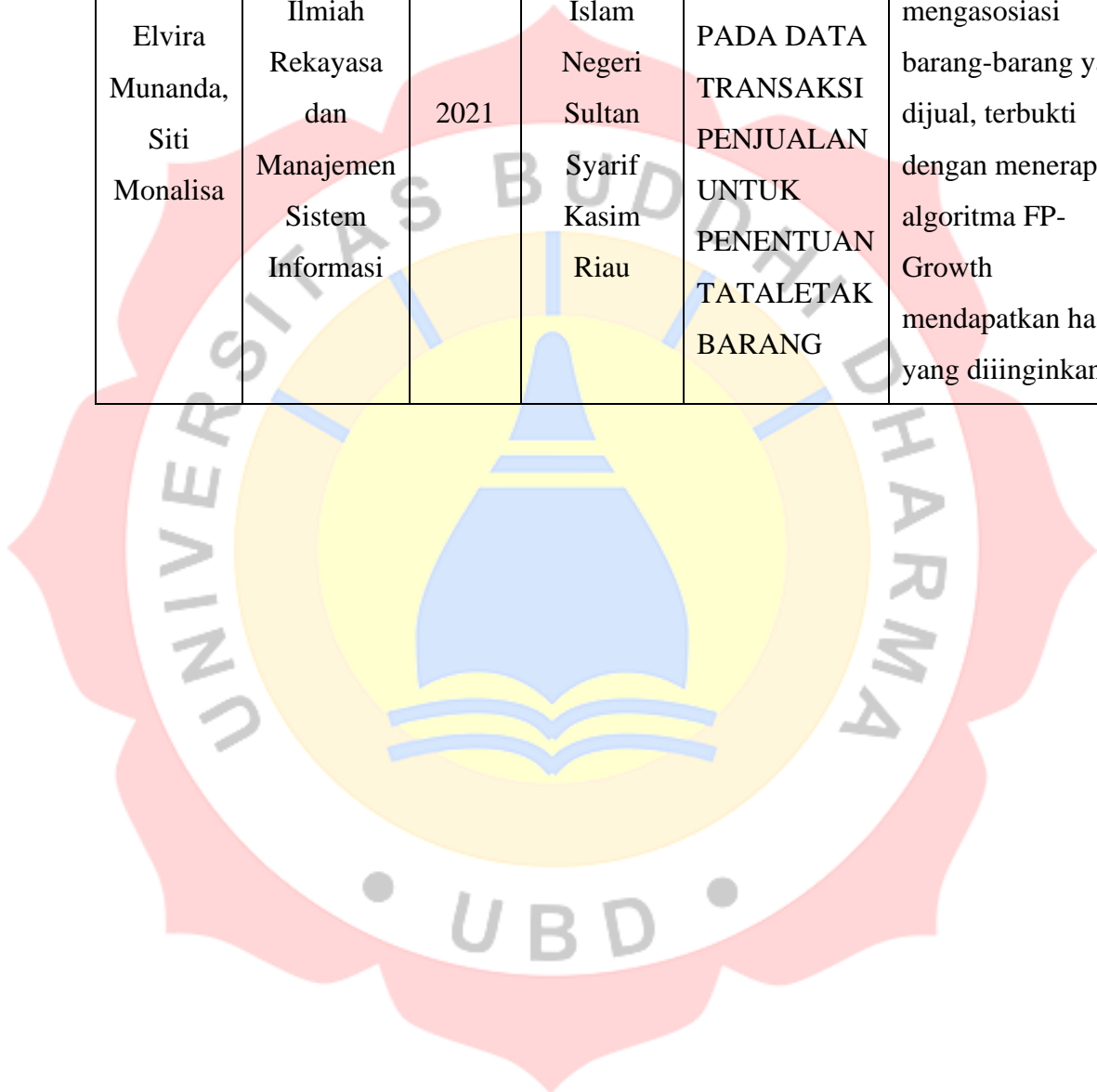


## 2.4.4 Tabel Perbandingan Jurnal

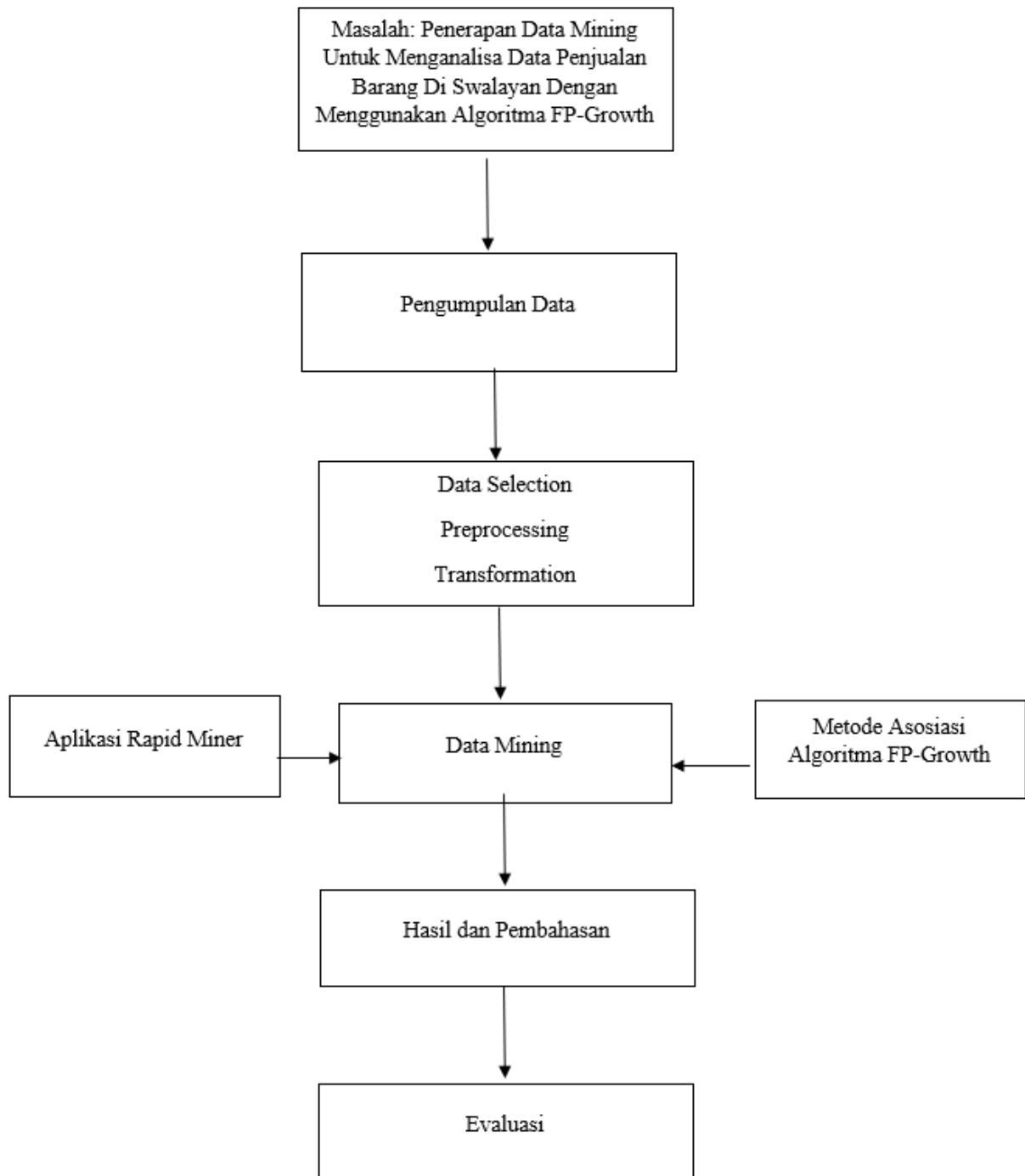
Peneliti	Nama Jurnal	Tahun	Institusi	Judul dan Metode yang digunakan	Kesimpulan
Alfannisa Annurullah Fajrin, Algifanri Maulana	Kumpulan Jurnal Ilmu Komputer (KLIK)	2020	Universitas Teknokrat Indonesia	PENERAPAN DATA MINING UNTUK ANALISIS POLA PEMBELIAN KONSUMEN DENGAN ALGORITMA FPGROWTH PADA DATA TRANSAKSI PENJUALAN SPARE PART MOTOR	Menerapkan metode Data Mining dengan algoritma FP-Growth kedalam aplikasi untuk analisis pola pembelian konsumen sangat bermanfaat bagi perusahaan tersebut, karena TB-Damar akan mengetahui spare part mana yang banyak dibeli dan membantu dalam pemesanan spare part pada kantor pusat. dari radikalisme, informasi palsu atau <i>hoax</i> , dan isu SARA di Indonesia.

Peneliti	Nama Jurnal	Tahun	Institusi	Judul dan Metode yang digunakan	Kesimpulan
Rizky Fitria Haya, Desy Ramadani	Jurnal Nasional Informatika dan Teknik Jaringan	2020	Universitas Samudra	Penerapan Data Mining dalam Analisis Pola Pembelian Minuman dan Makanan di Kantin SMAN 4 Langsa Menggunakan Algoritma FP-Growth	Penelitian ini berhasil mendapatkan tujuannya dengan menemukan barang mana saja yang keluar secara bersamaan dan kemudian meletakkan barang tersebut dipintu keluar agar lebih efisiensi dalam pengambilan barang dan pencarian barang.

<b>Peneliti</b>	<b>Nama Jurnal</b>	<b>Tahun</b>	<b>Institusi</b>	<b>Judul dan Metode yang digunakan</b>	<b>Kesimpulan</b>
Elvira Munanda, Siti Monalisa	Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi	2021	Universitas Islam Negeri Sultan Syarif Kasim Riau	PENERAPAN ALGORITMA FP-GROWTH PADA DATA TRANSAKSI PENJUALAN UNTUK PENENTUAN TATALETAK BARANG	Metode yang digunakan terbukti cocok untuk dalam mengasosiasi barang-barang yang dijual, terbukti dengan menerapkan algoritma FP-Growth mendapatkan hasil yang diinginkan.



## 2.5. Kerangka Pemikiran



**Gambar 2.1 Kerangka Pemikiran**

Pada kerangka pemikiran tersebut penulis menggunakan algoritma FP-Growth untuk menganalisa data penjualan barang di suatu swalayan, data yang digunakan merupakan data sekunder dan didapat melalui situs web kaggle.com, langkah yang dilakukan pertama adalah

memilah data agar data tersebut dapat digunakan dan dapat diolah, lalu dengan menggunakan

metode asosiasi dengan algoritma FP-Growth dan bantuan aplikasi Rapid Miner data mining pada data tersebut dilakukan, setelah melakukan proses perhitungan data mining didapatkan lah hasil serta pembahasannya yang dapat dievaluasi untuk memecahkan suatu masalah yang ada.



## BAB III

### ANALISA MASALAH

#### 3.1 Analisa Kebutuhan

Pada tahap ini dilakukan pengumpulan bahan terkait dengan data transaksi swalayan dan algoritma FP-Growth untuk mengelompokkan barang apa saja yang memiliki pengaruh besar satu sama lain. Data diperoleh melalui

<https://www.kaggle.com/datasets/heeraldedhia/groceries-dataset>, berikut merupakan dataset asli yang digunakan:



	A	B	C	D
1	Member_number, Date,itemDescription			
2	1808,21-07-2015,	tropical fruit		
3	2552,05-01-2015,	whole milk		
4	2300,19-09-2015,	pip fruit		
5	1187,12-12-2015,	other vegetables		
6	3037,01-02-2015,	whole milk		
7	4941,14-02-2015,	rolls/buns		
8	4501,08-05-2015,	other vegetables		
9	3803,23-12-2015,	pot plants		
10	2762,20-03-2015,	whole milk		
11	4119,12-02-2015,	tropical fruit		
12	1340,24-02-2015,	citrus fruit		
13	2193,14-04-2015,	beef		
14	1997,21-07-2015,	frankfurter		
15	4546,03-09-2015,	chicken		
16	4736,21-07-2015,	butter		
17	1959,30-03-2015,	fruit/vegetable juice		
18	1974,03-05-2015,	packaged fruit/vegetables		
19	2421,02-09-2015,	chocolate		
20	1513,03-08-2015,	specialty bar		
21	1905,07-07-2015,	other vegetables		
22	2810,08-09-2015,	butter milk		
23	2867,12-11-2015,	whole milk		
24	3962,18-09-2015,	tropical fruit		
25	1088,30-11-2015,	tropical fruit		
26	4976,17-07-2015,	bottled water		
27	4056,12-06-2015,	yogurt		
28	3611,13-02-2015,	sausage		
29	1420,14-01-2015,	other vegetables		

Gambar 3.1 Gambar Data Set Asli

Berikut merupakan keterangan data yang digunakan.

Tabel 3.1 Tabel Keterangan Data

Atribut	Keterangan	Variabel
Member_number	Atribut ini berisikan nomor unik konsumen, sebagai pengenal antara konsumen satu dengan yang lainnya,	Integer
Date	Atribut ini berisikan tanggal, bulan dan tahun Member_number yang melakukan pembelian barang.	Date
ItemDescription	Atribut ini berisikan deskripsi barang	Varchar

Lalu data set tersebut di data cleaning bertujuan untuk membersihkan data-data yang tidak diperlukan. Di dataset yang penulis gunakan tidak ditemukan data bernilai *null* atau data yang ambigu sehingga pada dataset “Groceries\_dataset.csv” ini, penulis hanya menghapus kolom yang tidak mengalami ketergantungan yaitu kolom **Date**, dengan gambaran sebagai berikut:

Member_number	itemDescription
1000	soda
1000	canned beer
1000	sausage
1000	sausage
1000	whole milk
1000	whole milk
1000	pickled vegetables
1000	misc. beverages
1000	semi-finished bread
1000	hygiene articles
1000	yogurt
1000	pastry
1000	salty snack
1001	frankfurter
1001	frankfurter
1001	beef
1001	sausage
1001	whole milk
1001	soda
1001	curd
1001	white bread
1001	whole milk
1001	soda
1001	whipped/sour cream
1001	rolls/buns
1002	tropical fruit

Gambar 3.2 Gambar Data Set Cleaning

Jadi penulis hanya menggunakan kolom **Member\_number** dan **itemDescription** saja pada dataset “Groceries\_dataset.csv”.

### 3.2 Konstruksi Algoritma dan Metode

#### Algoritma FP Growth

Dalam perhitungan manual algoritma FP-Growth data yang akan digunakan adalah sebagai berikut:



Tabel 3.2 Tabel Data Perhitungan Manual

Member Number	Item Description
1011	Yogurt, Whole Milk, Rolls/buns, Other Vegetables
1006	Whole Milk, Rolls/buns
1016	Rolls/buns
1032	Yogurt, Whole Milk, Soda, Other Vegetables
1008	Yogurt, Soda
1092	Yogurt, Soda, Rolls/buns, Other Vegetables
1141	Yogurt, whole milk
1133	Yogurt, Whole Milk, rolls/buns

Langkah-langkah perhitungan:

1. Menentukan Minimal Support

Minimal Support yang di gunakan adalah 3.

2. Mencari Nilai Support

Nilai support yang didapat adalah:

item	Support
yogurt	6
whole milk	5
soda	3
rolls/buns	5
other vegetables	3

Gambar 3.3 Gambar Tabel Item Support

## 3. Mencari Nilai Prioritas

item	Support	Prioritas	Singkatan
yogurt	6	1	Y
whole milk	5	3	WM
soda	3	5	S
rolls/buns	5	2	R/B
other vegetables	3	4	OV

Gambar 3.4 Gambar Nilai Prioritas Item

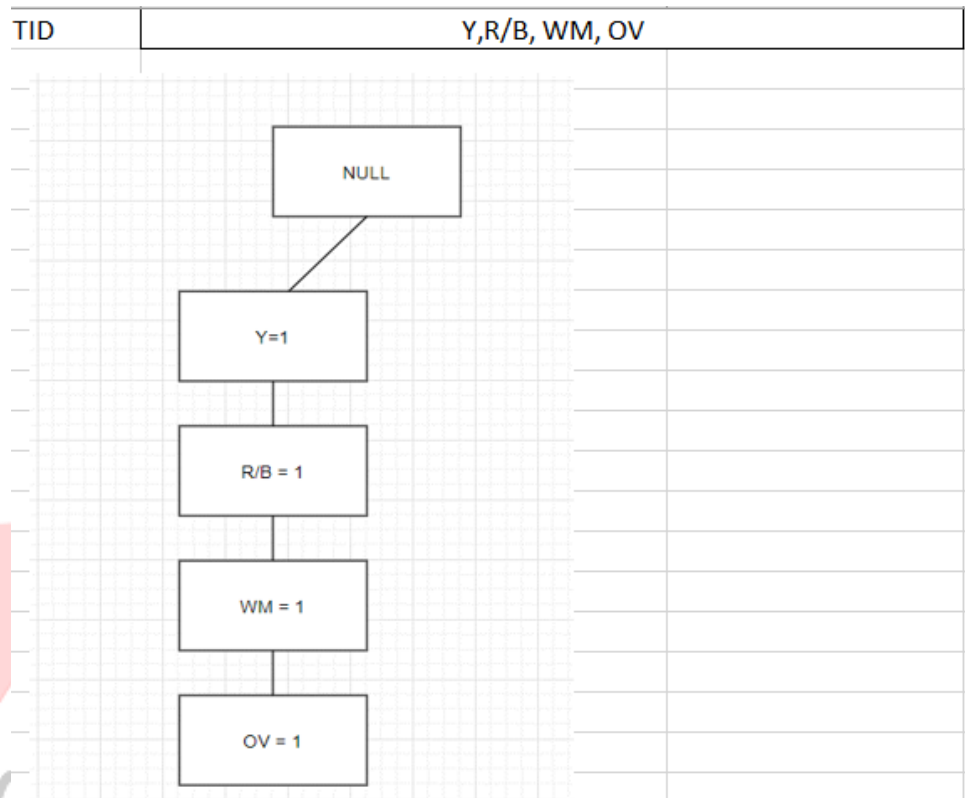
## 4. Mengurutkan berdasarkan prioritas

TID	Item	Urutan Item
1011	Yogurt, Whole Milk, Rolls/buns, Other Vegetables	Y,R/B, WM, OV
1006	Whole Milk, Rolls/buns	R/B, WM
1016	rolls/buns	R/B
1032	Yogurt, Whole Milk, Soda, Other Vegetables	Y,WM,OV,S
1008	Yogurt, Soda	Y,S
1092	Yogurt, Soda, Rolls/buns, Other Vegetables	Y,R/B,OV,S
1141	Yogurt, Whole milk	Y,WM
1133	Yogurt, whole milk, rolls/buns	Y,R/B,WM

Gambar 3.5 Gambar Urutan Item Berdasarkan Prioritas

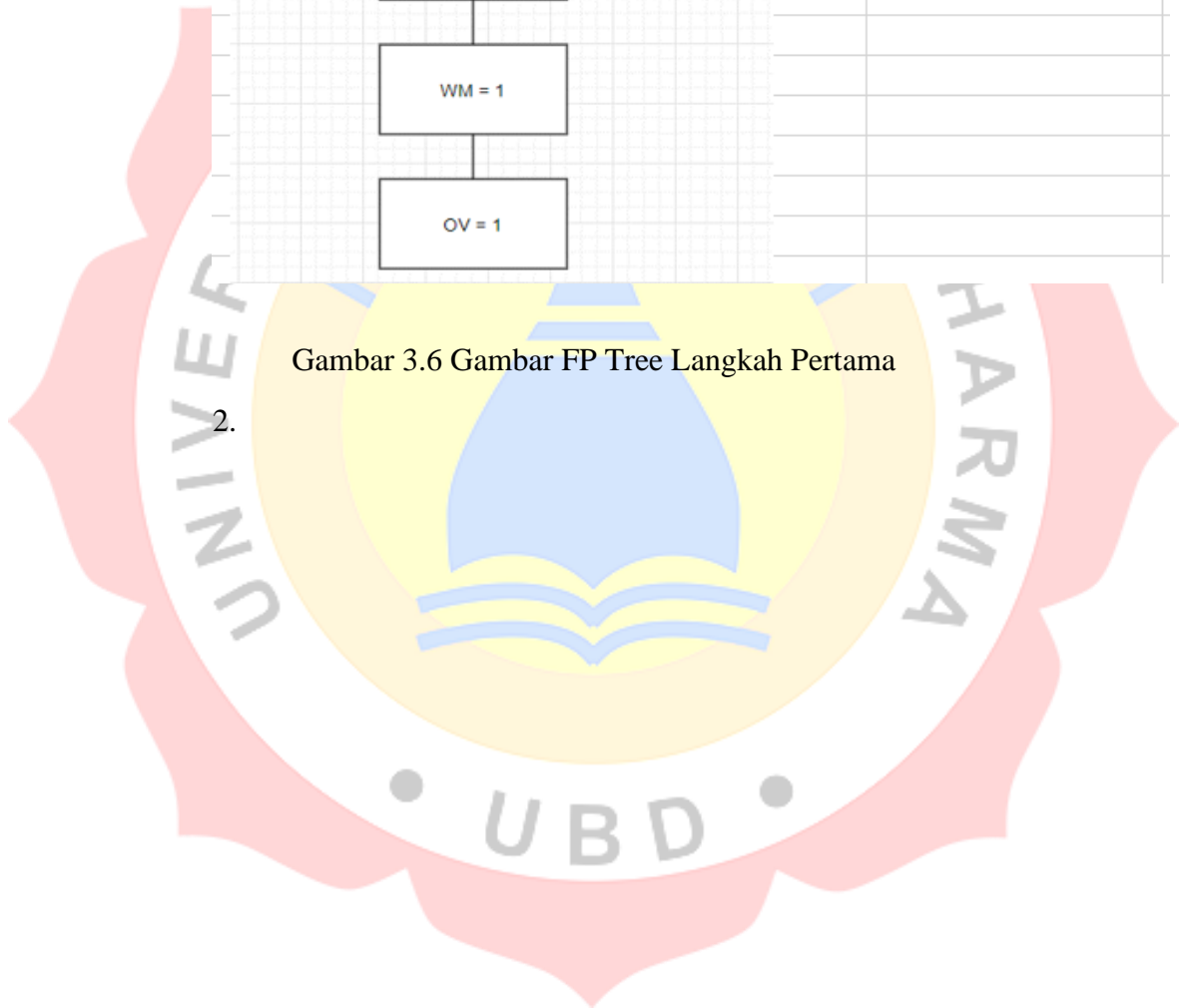
## 5. Membuat FP Tree dari data sebelumnya

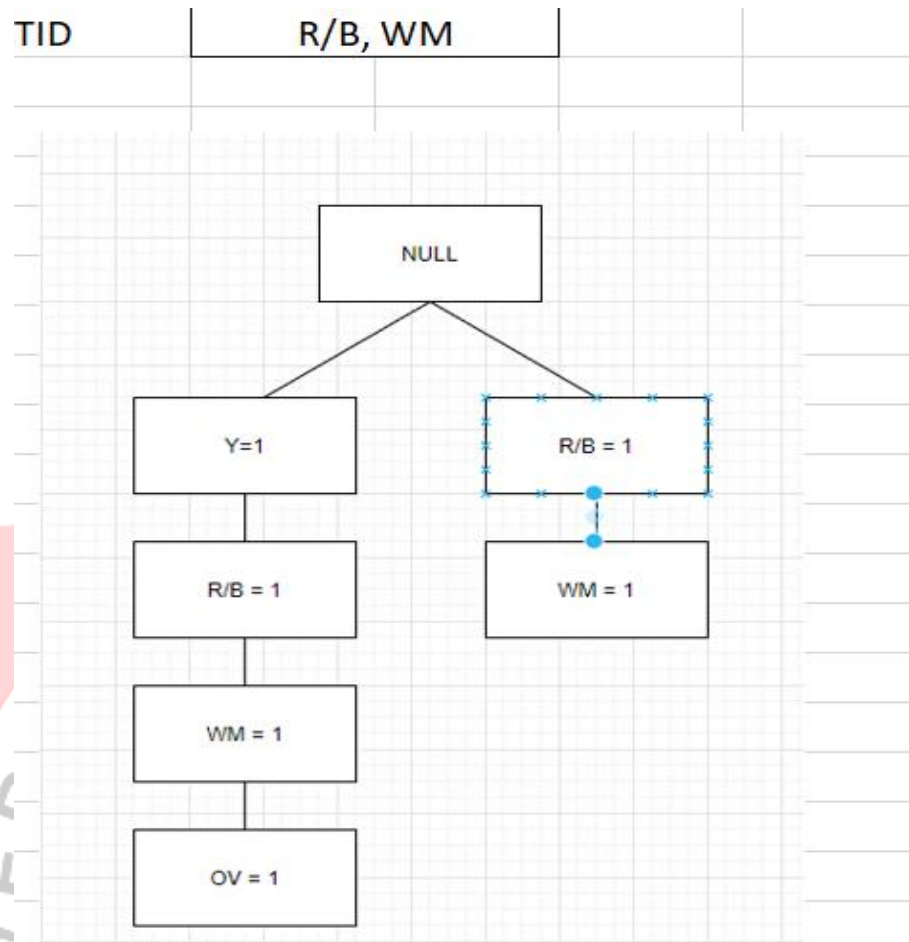
1.



Gambar 3.6 Gambar FP Tree Langkah Pertama

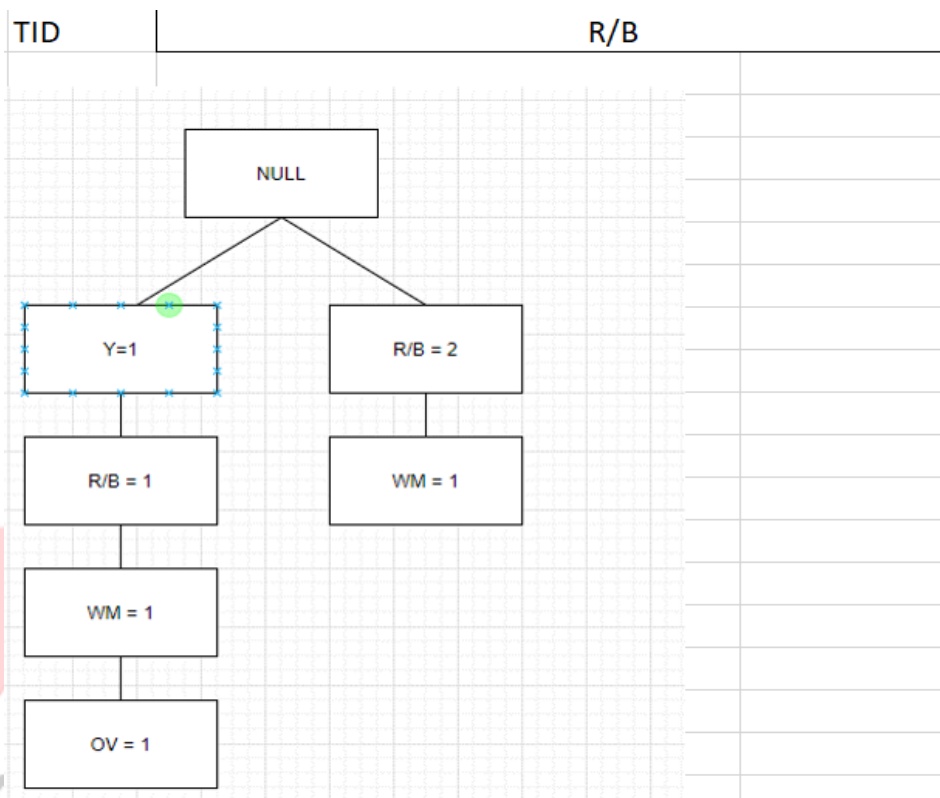
2.



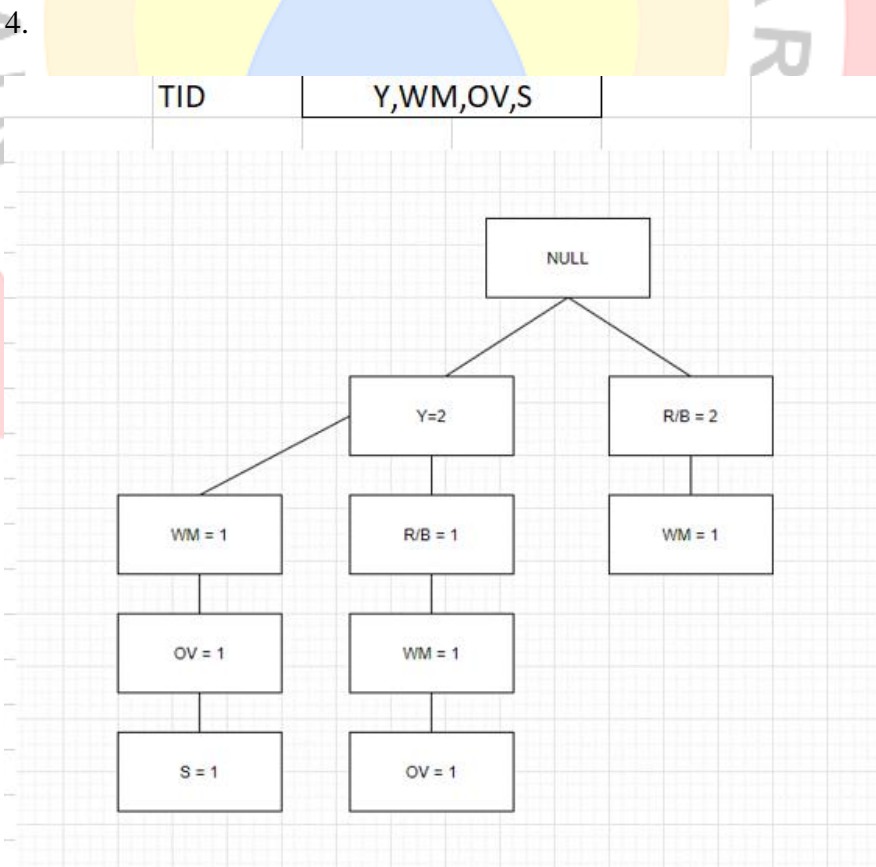


Gambar 3.7 Gambar FP Tree Langkah Kedua

3.

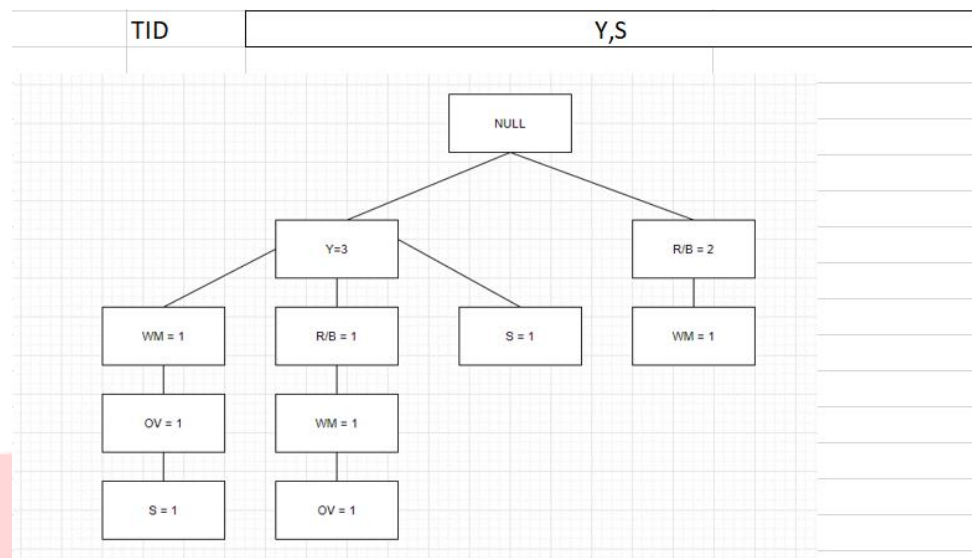


Gambar 3.8 Gambar FP Tree Langkah Ketiga



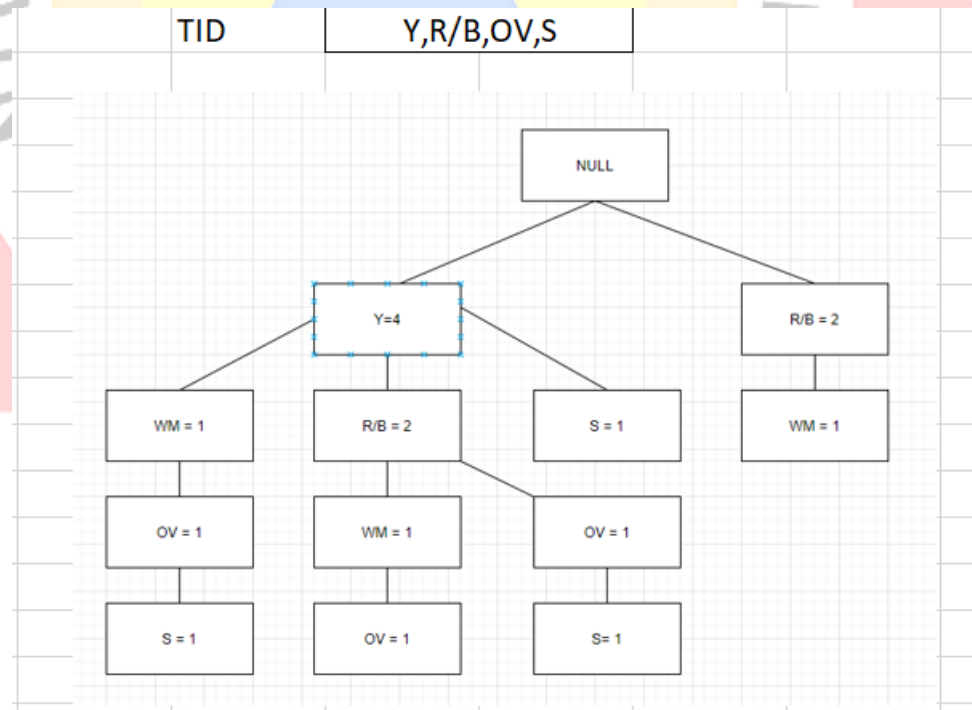
Gambar 3.9 Gambar FP Tree Langkah Keempat

5.



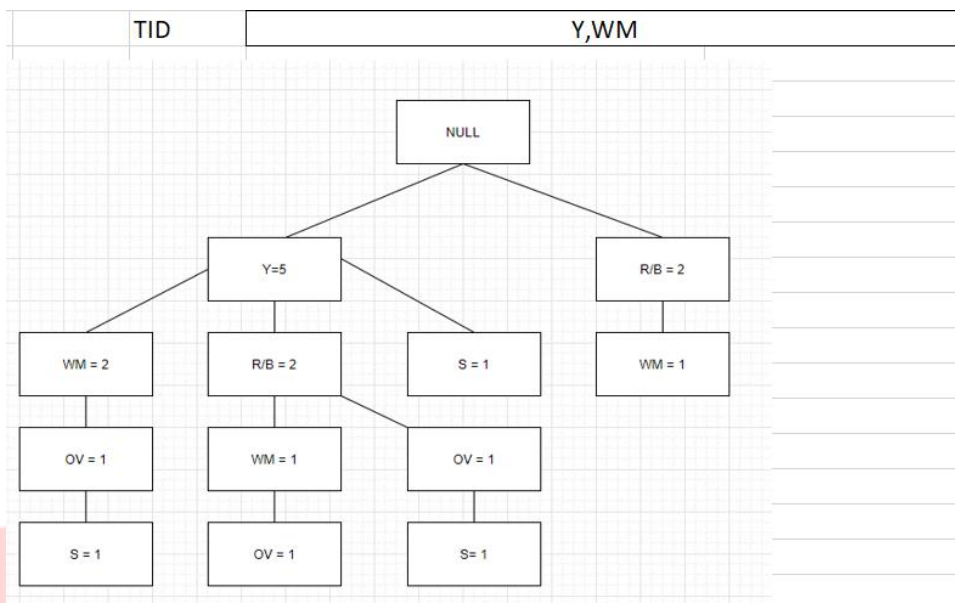
Gambar 3.10 Gambar FP Tree Langkah Kelima

6.

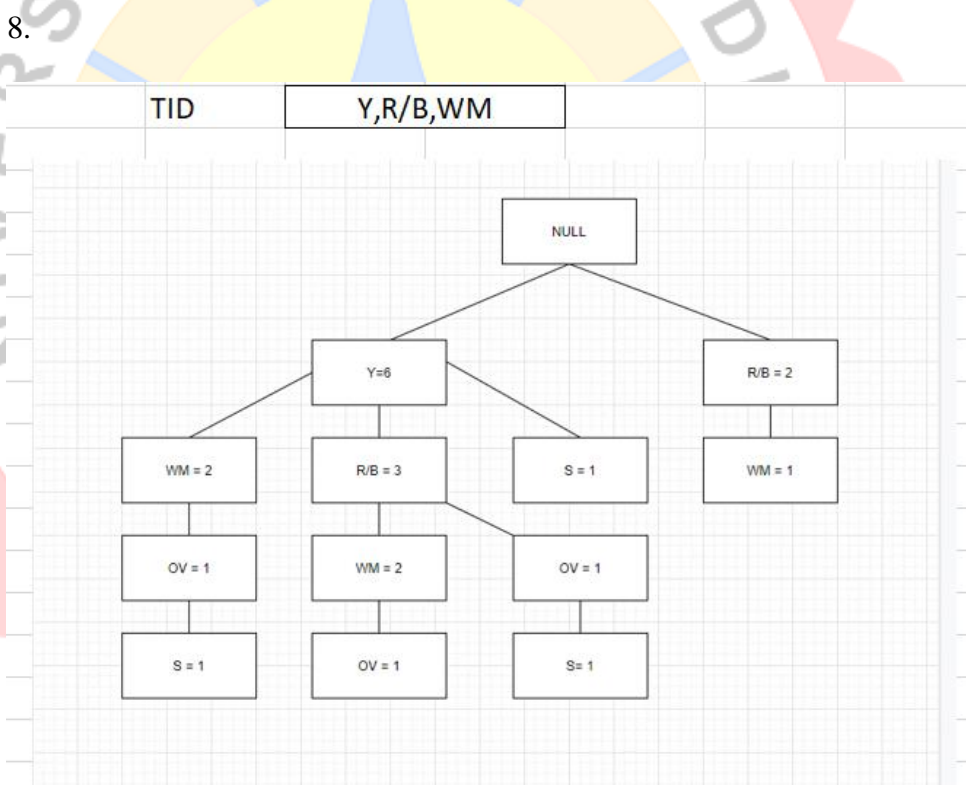


Gambar 3.11 Gambar FP Tree Langkah Keenam

7.



Gambar 3.12 Gambar FP Tree Langkah Ketujuh



Gambar 3.13 Gambar FP Tree Langkah Kedelapan

## 6. Membuat Frequent Pattern Generation

Dengan minimum support: 3

Tabel 3.3 Tabel Minimum Support Hitungan Manual

ITEM	Cond. Pat. Base	Cond. FP-Tree	FP-Generated
WM	(Y:2), (Y,R/B:2),(R/B:1)	Y:4 , R/B:3	YWM:4 , WMR/B :3
S	(Y, WM, OV:1),(Y:1), (Y, R/B , OV : 1)	Y:3	YS:3
R/B	(Y:3)	Y	YR/B:3
OV	(Y, WM : 1), (Y, R/B, WM : 1), (Y, R/B : 1)	Y:3	YOV:3

## 7. Mencari Support dan Confidence

Tabel 3.4 Tabel Hasil Support dan Confidence Perhitungan Manual

Itemset	Frequent	Support	Confidence
Y	6	0,75	1
WM	5	0,625	1
S	3	0,375	1
R/B	5	0,625	1
OV	3	0,375	1
Y,WM	4	0,5	0,666666667
WM, R/B	3	0,375	0,6
Y,S	3	0,375	0,5
Y,R/B	3	0,375	0,5
Y, OV	3	0,375	0,5

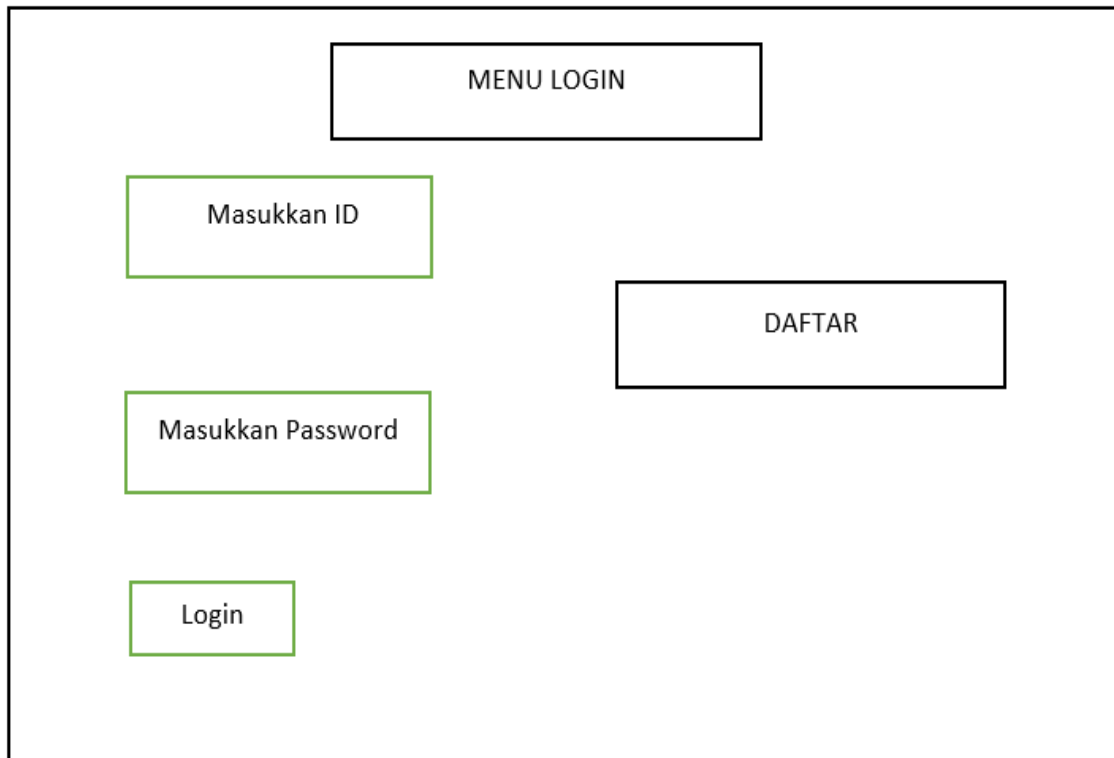
### 3.3 Perancangan Layar, Menu, Database

Pada bab ini akan menunjukkan tampilan perancangan layar pada *aplikasi* yang dikembangkan.

#### 3.3.1 Perancangan Halaman *Login*

Berikut merupakan tampilan perancangan awal pada halaman login, yang berisikan Username, password, tombol login dan tombol registrasi

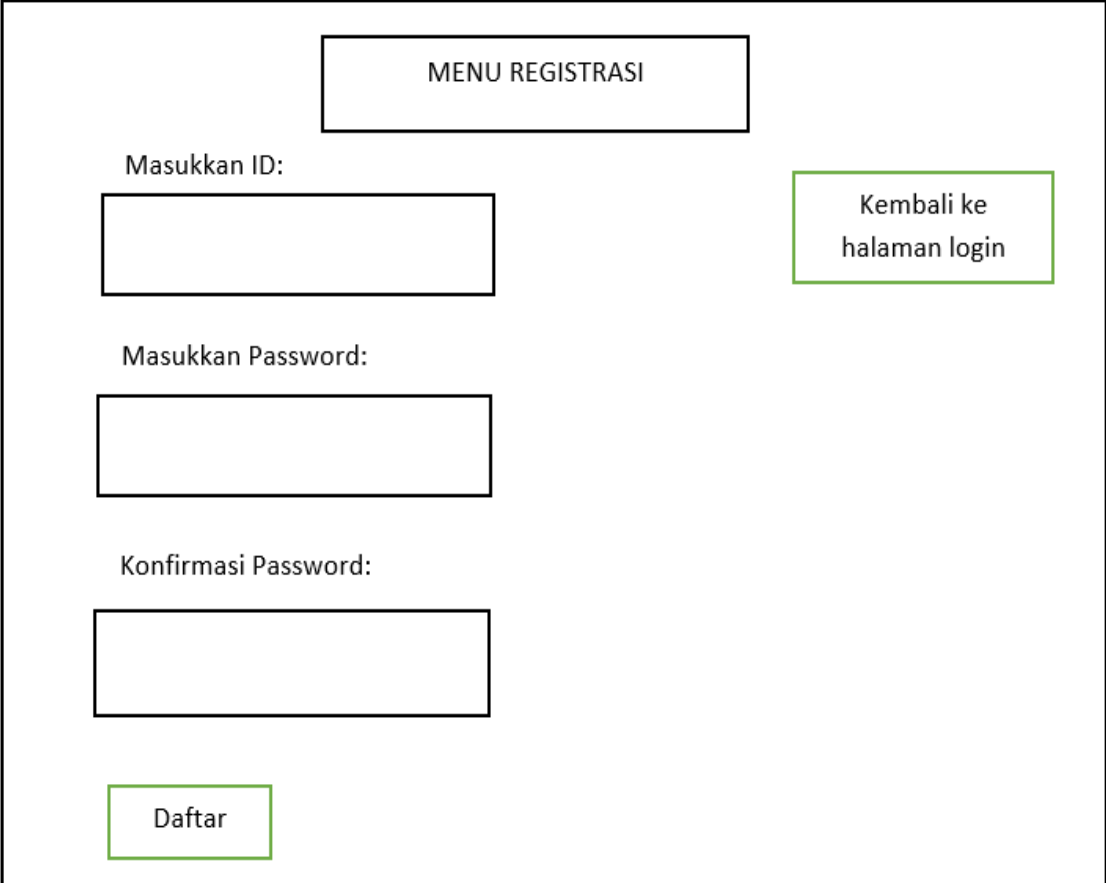




Gambar 3.14 Gambar Perancangan Halaman Login

### 3.3.2 Perancangan Halaman Registrasi

Berikut merupakan tampilan perancangan awal pada halaman registrasi yang berisikan Masukkan ID, Masukkan Password, Konfirmasi Password, tombol daftar dan tombol kembali ke menu login



MENU REGISTRASI

Masukkan ID:

Masukkan Password:

Konfirmasi Password:

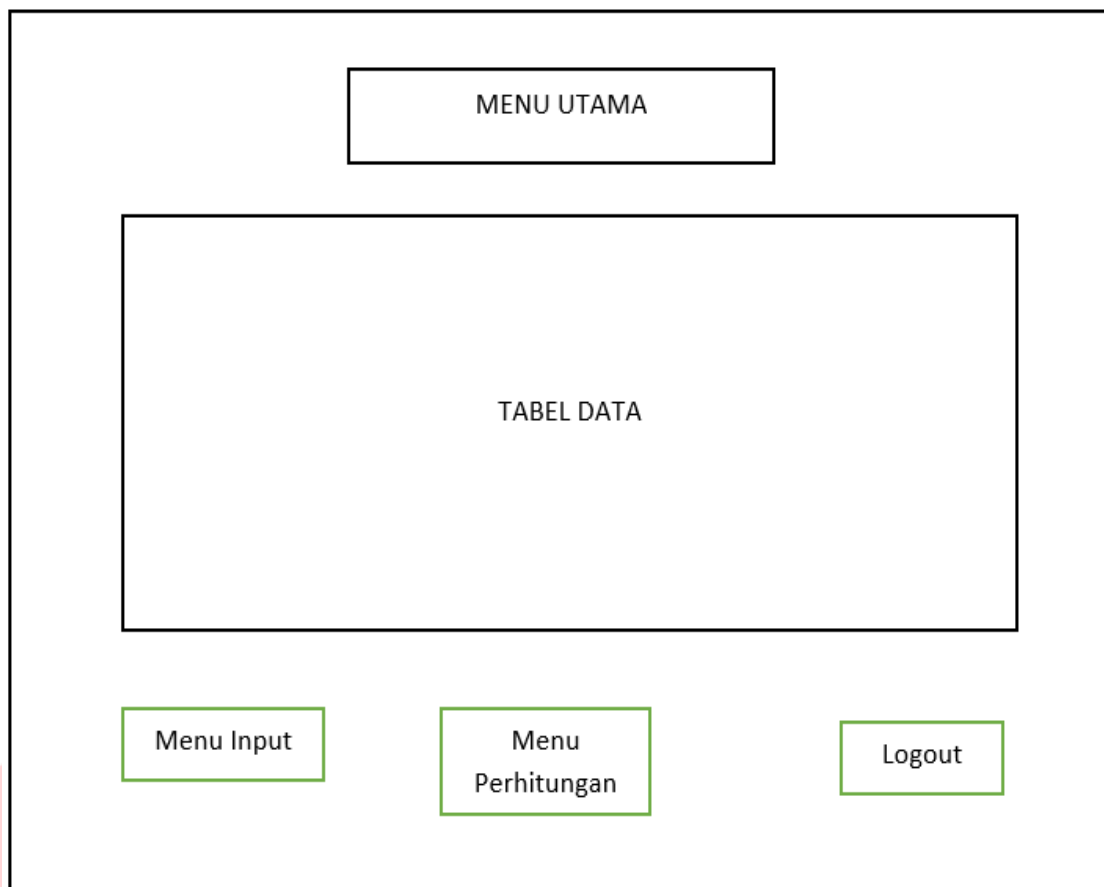
Daftar

Kembali ke halaman login

Gambar 3.15 Gambar Perancangan Halaman Registrasi

### 3.3.3 Perancangan Halaman Menu Utama

Berikut merupakan tampilan perancangan awal pada halaman menu utama yang berisikan tabel data, tombol menu input, tombol menu perhitungan dan tombol logout.



Gambar 3.16 Gambar Perancangan Halaman Menu Utama

### 3.3.4 Perancangan Halaman Menu Input Data

Berikut merupakan tampilan perancangan awal pada halaman input data yang berisikan member number, barang serta tombol tambah, tombol input data jika sudah selesai, tombol menu utama, tombol perhitungan dan tombol logout

MENU INPUT

Member Number:

Item:

Tambah Barang

Input Data

Menu Utama

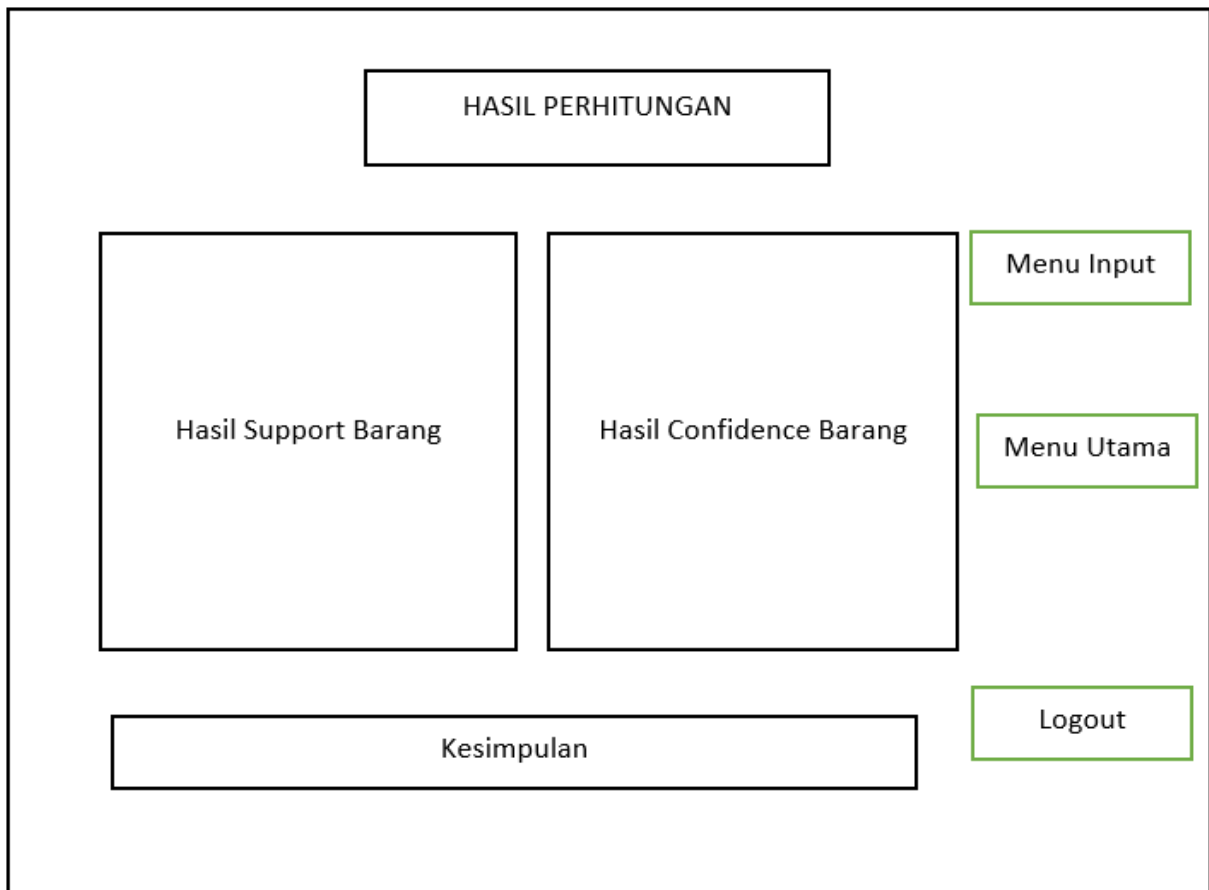
Menu Perhitungan

Logout

Gambar 3.17 Gambar Perancangan Halaman Menu Input

### 3.3.5 Perancangan Halaman Menu Perhitungan

Berikut merupakan tampilan perancangan awal pada halaman perhitungan terdapat hasil support dan hasil confidence dari barang yang diolah, kesimpulan, dan tombol-tombol menu lainnya seperti tombol menu input, tombol menu utama, dan logout



Gambar 3.18 Gambar Perancangan Halaman Menu Hitung

### 3.3.6 Perancangan Database

Pada perancangan database ini akan dijelaskan mengenai *database* yang dibuat dan digunakan dalam program yang akan dibuat. Berikut adalah rancangan *database* yang akan dibuat:

b. Tabel user

Tabel *accounts* digunakan untuk menyimpan data akun pengguna yang terdaftar ke dalam sistem. Tabel ini menyimpan data *username* dan *password*.

Tabel 3.5 Tabel User

<b>Nama Field</b>	<b>Type Field</b>
<i>username</i>	<i>Varchar (32)</i>
<i>password</i>	<i>Varchar (32)</i>

## c. Tabel Transaksi

Tabel transaksi berisi member number dan nama produk. Tabel transaksi digunakan untuk melihat transaksi yang terjadi pada suatu member terhadap barang yang di belinya.

Tabel 3.6 Tabel Transaksi

<b>Nama Field</b>	<b>Type Field</b>
Member number	Varchar (32)
Nama produk	Varchar (32)