

**PENERAPAN *DATA MINING* PENJUALAN RUMAH  
DENGAN MENGGUNAKAN *ALGORITMA APRIORI***

**SKRIPSI**



**Oleh:  
Frengki Manullang  
150210119**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KOMPUTER  
UNIVERSITAS PUTERA BATAM  
TAHUN 2020**

**PENERAPAN *DATA MINING* PENJUALAN RUMAH  
DENGAN MENGGUNAKAN *ALGORITMA APRIORI***

**SKRIPSI**

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



**Oleh:  
Frengki Manullang  
150210119**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK DAN KOMPUTER  
UNIVERSITAS PUTERA BATAM  
TAHUN 2020**

## PERNYATAAN

Yang bertanda tangan di bawah ini saya :

Nama : Frengki Manullang  
NPM : 150210119  
Fakultas : Teknik dan Komputer  
Program Studi : Teknik Informatika

Menyatakan bahwa “Skripsi” yang saya buat dengan judul :

**Penerapan *Data Mining* Penjualan Rumah dengan menggunakan *Algoritma Apriori***

Adalah hasil karya sendiri dan bukan “duplikasi” dari karya orang lain. Sepengetahuan saya, didalam 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 didalam naskah skripsi ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah skripsi ini digugurkan dan gelar 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, 17 Februari 2020



**Frengki Manullang**  
150210119

**PENERAPAN *DATA MINING* PENJUALAN RUMAH  
DENGAN MENGGUNAKAN *ALGORITMA APRIORI***

**SKRIPSI**

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

**Oleh:  
Frengki Manullang  
150210119**

**Telah disetujui oleh Pembimbing pada tanggal  
seperti tertera di bawah ini**

**Batam, 17 Februari 2020**

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

## ABSTRAK

Perkembangan teknologi informasi mulai memasuki bagian terpenting dalam suatu bidang pekerjaan, informasi yang di dapat dalam suatu pekerjaan menyangkut hal umum dan khusus. Komputer merupakan salah satu hasil dari perkembangan teknologi canggih yang dapat digunakan untuk membantu segala jenis pekerjaan termasuk transaksi pada pekerjaan bisnis dengan kemudahan yang ada didalam *software* komputer seorang user dapat melakukan operasional transaksi dengan cepat. Penggunaan komputer juga dibutuhkan oleh seorang marketing penjualan rumah. *PT. Glory Point* adalah perusahaan yang bergerak dibidang *Developer*, berkedudukan diwilayah kota Batam. Penjualan rumah pada perusahaan *PT. Glory Point* saat ini masih belum stabil dan mengakibatkan banyaknya rumah kosong dan tidak terawat karena tidak laku serta tempatnya yang jauh dari area transportasi umum, tempat pusat perbelanjaan, serta perkotaan. Jumlah unit perumahan pada *PT. Glory Point* berjumlah 236 unit diantaranya tipe 36/96, tipe 36/72b, tipe 42/78, tipe 46/84 dan tipe 36/72a yang terdiri dari 30 unit permasing-masing tipe, tipe 52/91 (23 unit), tipe 52/105 (20 unit), tipe 60/59 (20 unit), tipe 84 (12 unit), tipe 105 (11 unit). Dengan banyaknya data, maka dibutuhkan sebuah metode agar mempermudah hasil prediksi dalam penjualan rumah dengan menggunakan metode *algoritma apriori* dan bantuan *software tanagra*. *Algoritma apriori* termasuk jenis aturan asosiasi pada *data mining*. *Data mining* yang bersumber dari data komputer dan data transaksi penjualan kemudian diolah menggunakan metode *algoritma apriori*. Dengan perhitungan *algoritma apriori* maka diperoleh hasil penjualan rumah bertipe yang paling banyak dibeli oleh konsumen yaitu tipe 36/96, tipe 36/72b, tipe 42/78, tipe 46/84 dan tipe 36/72a dengan nilai *minimum support* 30% dan nilai *confidence* 90% dengan *finale rule* yang diperoleh sebanyak 12 *rules*.

**Kata Kunci:** *Data mining*, Penjualan rumah, *Algoritma Apriori*, *Software Tanagra*

## **ABSTRACT**

*The development of information technology began to enter the most important part in a field of work, the information obtained in a job concerning general and specific matters. The computer is one of the results of the development of technology used to help all transactions in business with the aim of facilitating and accelerating operations, one of the businesses that use computers is marketing home sales of PT. Glory Point. PT. Glory Point is a company engaged in the Developer, domiciled in the city of Batam. Home sales at PT. Glory Point is currently still unstable and results in many vacant and unkempt homes because they do not sell well and are far from urban areas and malls. The number of housing units at PT. Glory Points are 236 units including type 36/96, type 36 / 72b, type 42/78, type 46/84 and type 36 / 72a consisting of 30 units per type, 52/91 (25 units), type 52 / 105 (20 units), type 60/59 (18 units), type 84 (12 units), type 105 (11 units). With so much data, then we need a method in order to facilitate the results of predictions in home sales by using a priori algorithm method and the help of Tanagra software. A priori algorithm including the type of association rules in data mining. Data mining sourced from computer data and sales transaction data is then processed using a priori algorithm method. With the a priori algorithm calculation, the result of the sale of the type of houses most bought by consumers are type 36/96, type 36 / 72b, type 42/78, type 46/84 and type 36 / 72a with a minimum support value of 30% and the value of confidence 90% with 12 rule finale rules.*

**Keyword:** *Data mining, Home sale, Apriori algorithm, Tanagra software*

## KATA PENGANTAR

Puji dan syukur kepada Tuhan Yesus yang telah melimpahkan segala berkat 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, dorongan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

1. Rektor Universitas Putera Batam.
2. Bapak Andi Maslan, S.T., M.SI Ketua Program Studi Teknik Informatika Universitas Putera Batam.
3. Koko Handoko, S.Kom., M.Kom selaku pembimbing Skripsi pada Program Studi Teknik Informatika Universitas Putera Batam.
4. Ibu Nia Ekawati, S.Kom., M.SI selaku pembimbing akademik pada Program Studi Teknik Informatika Universitas Putera Batam selama perkuliahan .
5. Dosen dan Staff Universitas Putera Batam.
6. Erlin Winata selaku accounting di Perusahaan *PT. Glory Point*.
7. Kedua orang tua dan adik-adik yang selalu mendoakan.
8. Rekan-rekan mahasiswa/i Universitas Putera Batam yang turut memberikan doa dan dukungannya.

9. Meliani selaku sahabat seangkatan Teknik Informatika yang selalu ringan tangan membantu.
10. Sahabat Unit Kegiatan Mahasiswa Kristen Universitas Putera Batam.
11. Kelompok Tumbuh Bersama (KTB Daud) (Heriyenjuz, Yesaya, Yosua) yang selalu memberi saran, masukan dan mendoakan.
12. Kelompok Tumbuh Bersama (KTB Grace) (Devi, Indah, Yehesy, Strahel, Nelva, Nurhaida) yang selalu mendoakan.
13. Sahabat Persekutuan Mahasiswa Kristen Batam yang selalu mendoakan.
14. Sahabat Alumni Unit Kegiatan Mahasiswa Kristen Universitas Putera Batam yang selalu mendesak proses penulisan skripsi dan mendoakan.
15. Serta pihak-pihak lain yang tidak dapat disebutkan satu per satu.

Semoga Tuhan Yesus selalu memberkati dan melindungi kita semua menurut segala berkat karunia dan keselamatanNya, Amin.

Batam, 17 Februari 2020



(Frengki Manullang)

## DAFTAR ISI

Halaman

<b>HALAMAN SAMPUL</b> .....	<b>i</b>
<b>HALAMAN JUDUL</b> .....	<b>ii</b>
<b>PERNYATAAN</b> .....	<b>iii</b>
<b>HALAMAN PENGESAHAN</b> .....	<b>iv</b>
<b>ABSTRAK</b> .....	<b>v</b>
<b>ABSTRACT</b> .....	<b>vi</b>
<b>KATA PENGANTAR</b> .....	<b>vii</b>
<b>DAFTAR ISI</b> .....	<b>ix</b>
<b>DAFTAR GAMBAR</b> .....	<b>xi</b>
<b>DAFTAR TABEL</b> .....	<b>xii</b>
<b>DAFTAR RUMUS</b> .....	<b>xiii</b>
<b>BAB I PENDAHULUAN</b>	
1.1. Latar Belakang .....	1
1.2. Identifikasi Masalah .....	4
1.3. Batasan Masalah .....	4
1.4. Rumusan Masalah .....	5
1.5. Tujuan Penelitian .....	5
1.6. Manfaat Penelitian .....	6
1.6.1. Manfaat Teoritis .....	6
1.6.1. Manfaat Praktis .....	6
<b>BAB II LANDASAN TEORI</b>	
2.1. <i>Knowledge Discovery In Database (KDD)</i> .....	7
2.2. <i>Data Mining</i> .....	8
2.2.1. Proses Tahapan <i>Data Mining</i> .....	9
2.2.2. Pengelompokan <i>Data Mining</i> .....	10
2.2.3. Fungsi <i>Data Mining</i> .....	11
2.2.4. Kategori <i>Data Mining</i> .....	12
2.2.5. Manfaat <i>Data Mining</i> .....	12
2.2.6. Penerapan <i>Data Mining</i> .....	13
2.3. Metode <i>Data Mining</i> .....	14
2.3.1. <i>Algoritma Apriori</i> .....	14
2.3.2. Analisa Pola Frekuensi Tinggi dengan <i>Algoritma Apriori</i> .....	15
2.3.3. Pembentukan Aturan Asosiasi .....	15
2.4. <i>Software</i> Pendukung .....	16
2.5. Penelitian Terdahulu .....	17
2.6. Kerangka Pemikiran .....	23
<b>BAB III METODE PENELITIAN</b>	
3.1. Desain Penelitian .....	24
3.2. Metode Pengumpulan Data .....	26
3.2.1. Observasi .....	26
3.2.2. Wawancara .....	26
3.2.3. Studi Pustaka .....	26

3.3. Operasional Variabel .....	27
3.4. Metode Perancangan Sistem .....	29
3.5. Lokasi dan Jadwal Penelitian .....	30
3.5.1. Lokasi Penelitian .....	30
3.5.2. Jadwal Penelitian .....	30

**BAB IV HASIL DAN PEMBAHASAN**

4.1. Hasil Penelitian .....	31
4.1.1. Analisa Data .....	31
4.1.2. Membuat Tabel Tabulasi Penjualan .....	33
4.1.3. Menghitung Pola Frekuensi 1 <i>Itemset</i> .....	36
4.1.4. Menghitung Kombinasi 2 <i>Itemset</i> .....	39
4.1.5. Menghitung Kombinasi 3 <i>Itemset</i> .....	44
4.1.6. Menghitung Kombinasi 4 <i>Itemset</i> .....	54
4.1.7. Menghitung Kombinasi 5 <i>Itemset</i> .....	58
4.1.8. Pembentukan Pola Aturan <i>Association</i> .....	59
4.2. Pengujian aplikasi <i>Tanagra</i> .....	65
4.3. Pembahasan .....	72

**BAB V SIMPULAN DAN SARAN**

5.1. Simpulan .....	75
5.2. Saran .....	76

**DAFTAR PUSTAKA**

**LAMPIRAN**

- Lampiran 1. Pendukung Penelitian
- Lampiran 2. Daftar Riwayat Hidup
- Lampiran 3. Surat Keterangan Penelitian

## DAFTAR GAMBAR

Halaman

<b>Gambar 2. 1</b> Tahapan <i>Knowledge Discovery In Database</i> (KDD).....	9
<b>Gambar 2. 2</b> <i>Tanagra Software</i> .....	16
<b>Gambar 2. 3</b> Kerangka Pemikiran.....	23
<b>Gambar 3. 1</b> Desain Penelitian.....	24
<b>Gambar 3. 2</b> Metode Perancangan Sistem.....	29
<b>Gambar 4. 1</b> Format tubular pada lembar kerja <i>Microsoft excel</i> .....	65
<b>Gambar 4. 2</b> Tabel Tabular dalam <i>tab Add-ins</i> .....	66
<b>Gambar 4. 3</b> Tampilan <i>database</i> pada aplikasi <i>Tanagra</i> .....	66
<b>Gambar 4. 4</b> Tampilan <i>database</i> yang ada pada <i>Tanagra</i> .....	67
<b>Gambar 4. 5</b> Tampilan item-item dari <i>database</i> .....	67
<b>Gambar 4. 6</b> Tampilan Memasukkan item-item dari atribut ke <i>input</i> .....	68
<b>Gambar 4. 7</b> Tampilan setelah item dari atribut dipindahkan ke <i>input</i> .....	68
<b>Gambar 4. 8</b> Menentukan <i>minimum support</i> 0,30 dengan <i>max. support</i> 1.....	69
<b>Gambar 4. 9</b> Nilai <i>Frequenst Itemset</i> Data. ....	70
<b>Gambar 4. 10</b> Tampilan bagian <i>Apriori</i> .....	70
<b>Gambar 4. 11</b> Menentukan nilai <i>confidence</i> .....	71
<b>Gambar 4. 12</b> Menampilkan nilai <i>confidence itemset</i> .....	72
<b>Gambar 4. 13</b> Hasil nilai <i>confidence itemset</i> .....	72

## DAFTAR TABEL

Halaman

<b>Tabel 3. 1</b> Operasional Variabel .....	28
<b>Tabel 3. 2</b> Jadwal Penelitian .....	30
<b>Tabel 4. 1</b> Data Transaksi Penjualan.....	32
<b>Tabel 4. 2</b> Tabel Tabulasi Penjualan .....	34
<b>Tabel 4. 3</b> Kode Tipe Rumah.....	36
<b>Tabel 4. 4</b> Pola frekuensi 1 <i>itemset</i> .....	37
<b>Tabel 4. 5</b> Hasil frekuensi 1 <i>itemset</i> .....	38
<b>Tabel 4. 6</b> Perhitungan kombinasi 2 <i>itemset</i> .....	39
<b>Tabel 4. 7</b> Hasil Perhitungan Kombinasi 2 <i>itemset</i> .....	43
<b>Tabel 4. 8</b> Perhitungan kombinasi 3 <i>itemset</i> .....	44
<b>Tabel 4. 9</b> Hasil Perhitungan Kombinasi 3 <i>itemset</i> .....	53
<b>Tabel 4. 10</b> Perhitungan kombinasi 4 <i>itemset</i> .....	54
<b>Tabel 4. 11</b> Hasil Perhitungan Kombinasi 4 <i>itemset</i> .....	58
<b>Tabel 4. 12</b> Perhitungan kombinasi 5 <i>itemset</i> .....	58
<b>Tabel 4. 13</b> Hasil Perhitungan 2, 3, 4, dan 5 <i>itemset</i> .....	59
<b>Tabel 4. 14</b> Hasil pola aturan <i>Association</i> .....	60
<b>Tabel 4. 15</b> Hasil aturan <i>Association</i> .....	63

## DAFTAR RUMUS

	Halaman
<b>Rumus 2. 1</b> Frekuensi tinggi 1 <i>Itemset</i> .....	15
<b>Rumus 2. 2</b> Menghitung kombinasi 2 <i>Itemset</i> .....	15
<b>Rumus 2. 3</b> Menghitung nilai <i>Confidence</i> .....	15

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Perkembangan teknologi informasi mulai memasuki bagian terpenting dalam suatu bidang pekerjaan, informasi yang di dapat dalam suatu pekerjaan menyangkut hal umum dan khusus. Hal tersebut dapat berupa kumpulan data yang banyak, dikumpulkan kemudian diolah menjadi suatu informasi baru yang lebih baik, sederhana dan menarik. Data ini di dapat dari setiap transaksi yang biasanya kita lakukan setiap hari didalam kebanyakan jenis pekerjaan. Salah satunya pekerjaan yang menghasilkan data cukup banyak adalah data penjualan. Penjualan tersebut dapat berupa jasa maupun produk. Penjualan jasa erat kaitannya dengan keahlian, waktu, hasil dan harga yang kita tawarkan kepada konsumen. Sedangkan penjualan produk erat kaitannya dengan tipe, merek, model, fungsi dan harga produk tersebut.

Dalam dunia bisnis khususnya di industri atau perusahaan tidak sedikit pesaing lama dan baru yang menjadi tantangan bagi seorang pemilik produk untuk memasarkan produknya. Dalam hal ini pemilik produk sangat dituntut untuk mengembangkan atau menemukan inovasi terbaru guna untuk meningkatkan penjualan suatu produk kepada pembeli dan konsumen, salah satunya dengan cara meningkatkan manfaat sebuah data penjualan yang ada. Penjualan yang ada dilakukan setiap harinya sehingga membuat data itu bertambah banyak, yang semakin lama semakin banyak. Data yang dikumpulkan berfungsi sebagai laporan penjualan produk perusahaan, banyaknya data tersebut berguna untuk di olah dan di analisa untuk menjadi informasi yang dapat berguna.

Rumah merupakan salah satu tempat tinggal yang dihuni oleh setiap manusia. Rumah adalah produk yang ditawarkan kepada masyarakat dengan menjadikan itu kebutuhan yang sangat penting didalam kehidupan. Rumah terus dibuat dengan bentuk dan model yang sangat ideal sesuai perkembangan zaman. Rumah dengan bentuk yang ideal dan terjangkau dari segi harga dan fasilitas menjadi idaman setiap orang untuk membelinya. Pertimbangan dalam pembelian rumah yang dilakukan konsumen menjadi tolak ukur dalam memilih rumah idaman. Melalui koran, brosur, media sosial *facebook* dan baliho, pemasaran dilakukan oleh penjual rumah dengan tujuan memikat hati konsumen.

Penjualan rumah pada salah satu *developer* perusahaan saat ini masih belum stabil dikarenakan belum begitu dilirik secara langsung dimasyarakat kota Batam. Kualitas rumah yang murah dan strategis mempengaruhi minat masyarakat dalam menentukan kesukaan terhadap rumah yang akan dibeli, apabila rumah dalam jumlah besar tidak terjual maka penjualan tidak stabil maka mengakibatkan banyak rumah akhirnya menjadi rumah kosong dan tidak terawat dengan baik. Penataan dan pemeliharaan rumah tersebut juga menjadi terbengkalai. Faktor lingkungan merupakan salah satu faktor yang dipertimbangkan konsumen sebelum membeli rumah. (Anastasia, 2013).

Kurangnya Minat konsumen membeli suatu rumah dikarenakan lokasi yang kurang strategis, jauh dari transportasi umum, tempat perbelanjaan dan sekolah. Pembangunan dan pemasaran yang dibangun oleh *developer* beralamat di Tanjung Riau dan Tiban Koperasi. Dengan kondisi sarana transportasi umum yang belum banyak beroperasi melewati daerah tersebut dan jangkauan sekolah anak serta pusat

perbelanjaan kemudian masyarakat mempertimbangkannya kembali.. Harga, tipe dan luas tanah yang tidak sesuai mengakibatkan konsumen secara tidak langsung juga mempertimbangkannya untuk membeli rumah. Dalam pengelolaan hasil penjualan rumah saat ini, seorang marketing menggunakan perangkat komputer dan mencatat data transaksi dengan tujuan mendapat data sebagai laporan akhir mereka kepada perusahaan. Namun data tersebut belum dapat digunakan sebagai data yang *valid* dalam memprediksi hasil penjualan rumah dimasa mendatang. Data transaksi pembelian rumah dengan cara konvensional masih diperuntukkan untuk mengetahui minat konsumen, perlu adanya cara untuk memprediksi hasil penjualan rumah untuk mengetahui minat konsumen terhadap rumah yang diminati berdasarkan hasil transaksi sebelumnya, sehingga persediaan rumah dapat memenuhi permintaan konsumen.

Data yang diperoleh dalam penelitian ini mencakup tipe, harga, luas tanah serta total penjualan perbulannya dalam tiga tahun terakhir. Data yang tersimpan dalam komputer dengan jumlah yang banyak ini digunakan sebagai *data mining*. *Data mining* adalah aturan pencarian pola atau informasi menarik dalam suatu kumpulan data terpilih dengan menggunakan metode tertentu (Nurul Adha, Lince Tomoria Sianturi, 2017).

*Data mining* yang bersumber dari data komputer dan transaksi penjualan kemudian di olah menggunakan *algoritma apriori*. *Algoritma apriori* adalah jenis aturan asosiasi pada *data mining*. *Apriori* adalah *algoritma* kelas yang membantu mempelajari peraturan asosiasi. *Algoritma* mencoba untuk menemukan himpunan bagian yang umum dalam kumpulan data (Ristianingrum, 2017)

## 1.2 Identifikasi Masalah

Berdasarkan latar belakang masalah diatas, maka identifikasi masalah adalah sebagai berikut.

1. Penjualan yang tidak stabil mengakibatkan banyak rumah akhirnya menjadi rumah kosong dan tak terawat.
2. Kurangnya minat konsumen membeli rumah dikarenakan lokasi yang tidak strategis, jauh dari transportasi umum, tempat perbelanjaan dan sekolah.
3. Harga, luas tanah dan tipe yang tidak sesuai mengakibatkan konsumen secara tidak langsung enggan untuk membeli.

## 1.3 Batasan Masalah

Untuk menghindari adanya penyimpangan dari judul dan tujuan yang sebenarnya, maka adapun batasan masalah yang ditemukan dalam penelitian ini sebagai berikut :

1. Lokasi Penelitian ini bertempat di *PT. Glory Point*.
2. Data yang digunakan dalam penelitian hanya tiga tahun terakhir Oktober 2016 sampai dengan September 2019.
3. Penelitian ini mengguakan *Algoritma Apriori* dan *Software Tanagra*.

#### 1.4 Rumusan Masalah

Berdasarkan uraian dari latar belakang, maka dapat dirumuskan suatu masalah penelitian yang akan dibahas yaitu:

1. Bagaimana *data mining* menggunakan *algoritma apriori* dalam mencari pola rumah yang diminati konsumen ?
2. Bagaimana lingkungan pembangunan rumah yang sudah ada kemudian dikembangkan guna untuk menarik minat konsumen ?
3. Bagaimana *software Tanagra* dapat membantu mengolah data perhitungan manual sehingga dapat mempermudah penjualan rumah ?

#### 1.5 Tujuan Penelitian

Berdasarkan uraian dari rumusan masalah, maka tujuan dalam melakukan penelitian ini adalah:

1. Untuk menemukan pola rumah yang diminati konsumen dibutuhkan strategi penjualan rumah yang lebih mengutamakan kebutuhan dan keinginan yang dicari konsumen.
2. Untuk mengembangkan pembangunan rumah kedepannya diperlukan pemilihan lingkungan yang strategis, berkualitas, murah dan mudah terjangkau dari segi transportasi, tempat perbelanjaan dan sekolah.
3. Untuk membandingkan perhitungan data yang dilakukan secara manual, terhadap perhitungan yang secara otomatis dengan penggunaan *software tanagra* sehingga mempermudah data penjualan.

## 1.6 Manfaat Penelitian

Dalam pelaksanaan penelitian ini, diharapkan dapat memberi manfaat antara lain :

### 1.6.1 Manfaat Teoritis

Adapun manfaat teoritis yang peneliti dapatkan dari penelitian ini adalah sebagai berikut :

1. Untuk menerapkan dan mengembangkan ilmu yang telah dipelajari selama perkuliahan dalam jurusan teknik informatika.
2. Tambahan *referensi* bagi pembaca yang akan melakukan penelitian lebih lanjut khusus bidang teknik informatika tentang *data mining* dalam memprediksi hasil penjualan menggunakan *algoritma apriori*.

### 1.6.2 Manfaat Praktis

Memperoleh pengetahuan baru tentang *data mining* dengan *algoritma apriori* dalam memprediksi hasil penjualan produk yang tepat dan dapat dijadikan *referensi* bagi perusahaan.

1. Dengan menggunakan metode ini, *developer* dan marketing pemasaran rumah kedepannya lebih mengutamakan keinginan dan kebutuhan konsumen.
2. Memperoleh pengetahuan baru tentang *data mining* metode *algoritma apriori* dalam memprediksi hasil penjualan rumah yang tepat dan dapat dijadikan referensi bagi perusahaan.

## BAB II

### LANDASAN TEORI

#### 2.1 *Knowledge Discovery In Database (KDD)*

Menurut (Fajhrin, 2018), *Knowledge Discovery In Database (KDD)* diartikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data. Proses KDD melibatkan hasil proses dari *data mining* (Proses mengekstrak kecendrungan pola suatu data), kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami. KDD sendiri diartikan sebagai keseluruhan proses *non-trivial* untuk mencari dan mengidentifikasi pola (*pattern*) dalam data, dimana pola yang ditemukan bersifat sah, baru, dapat bermanfaat dan dimengerti.

Menurut (Kennedi Tampubolon, hoga saragih, 2013), ada enam elemen yang paling esensial dalam teknik pencarian informasi atau pengetahuan dalam KDD yaitu :

1. Mengerjakan sejumlah data besar.
2. Diperlukan efisiensi berkaitan dengan volume data.
3. Mengutamakan ketetapan/keakuratan.
4. Membutuhkan pemakaian Bahasa tingkat tinggi.
5. Menggunakan beberapa bentuk dari pembelajaran otomatis.
6. Menghasilkan hasil yang menarik.

## 2.2 *Data Mining*

Menurut (Kennedi Tampubolon, hoga saragih, 2013) *data mining* merupakan proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakit dari berbagai *database* besar/ *data warehouse*.

Menurut (Fajhrin, 2018), *data mining* bukanlah suatu bidang yang sama sekali baru. Salah satu kesulitan untuk mendefinisikan *data mining* adalah kenyataan bahwa *data mining* mewarisi banyak aspek dan teknik dari bidang-bidang ilmu yang sudah mapan terebih dahulu. Berawal dari beberapa disiplin ilmu, *data mining* bertujuan untuk memperbaiki teknik tradisional sehingga bisa menangani :

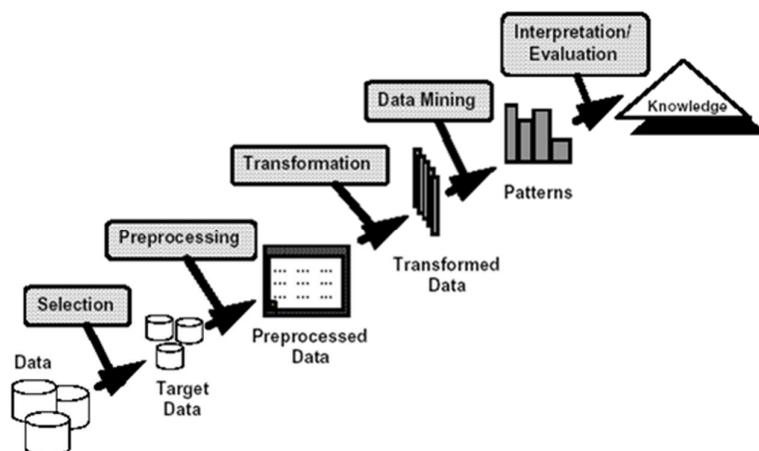
1. Jumlah data yang sangat besar.
2. Dimensi data yang tinggi.
3. Data yang heterogen dan berbeda sifat.

Menurut (Koko Handoko, 2019) *Data Mining* merupakan sebuah aktivitas yang dilakukan dengan menggabungkan berbagai ilmu statistika, matematika, *artificial intelligence* dan *machine learning* dengan tujuan untuk melakukan ekstraksi dan identifikasi dari berbagai informasi yang ada sehingga menghasilkan suatu pengetahuan baru yang bermanfaat pada hasil kedepannya pada sebuah *database* yang besar.

### 2.2.1 Proses Tahapan *Data Mining*

Menurut (Vulandari, 2017), *Data mining* merupakan salah satu dari rangkaian KDD. KDD berhubungan dengan teknik *integrasi* dan penemuan ilmiah, interpretasi dan visualisasi dari pola-pola sejumlah data. Serangkaian proses tersebut memiliki tahapan sebagai berikut :

1. Pembersihan data (untuk membuang data yang tidak konsisten dan *noise*).
2. Integrasi data (penggabungan data dari beberapa sumber).
3. Transformasi data (data diubah menjadi bentuk yang sesuai untuk di-*mining*).
4. Aplikasi teknik *Data Mining*, proses ekstraksi pola dari data yang ada).
5. Evaluasi pola yang ditemukan (proses interpretasi pola menjadi pengetahuan yang dapat digunakan untuk mendukung pengambilan keputusan).
6. Presentasi pengetahuan (dengan teknik visualisasi).



**Gambar 2. 1** Tahapan *Knowledge Discovery In Database (KDD)*

**Sumber:** Data Mining Teori dan Aplikasi Rapidminer

### 2.2.2 Pengelompokan *Data Mining*

Menurut (Fajhrin, 2018), pengelompokan *data mining* dibagi menjadi beberapa kelompok yaitu:

1. Deskripsi

Deskripsi merupakan cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data yang dimiliki.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih kearah numerik dari pada kearah kategori. Model yang dibangun menggunakan record lengkap yang menyediakan nilai variabel target sebagai nilai prediksi.

3. Prediksi

Prediksi menerka sebuah nilai yang belum diketahui dan juga memperkirakan nilai untuk masa mendatang.

4. Klasifikasi

Dalam klasifikasi terdapat target variabel kategori, misal penggolongan pendapatan dapat dipisahkan dalam tiga kategori yaitu tinggi, sedang, dan rendah.

5. Pengklusteran

Merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan.

## 6. Asosiasi

Assosiasi bertugas menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

### 2.2.3 Fungsi *Data Mining*

Menurut (Vulandari, 2017), fungsi - fungsi yang umum diterapkan dalam *data mining* (haskett, 2000) yaitu :

1. *Assosiation*, adalah proses untuk menemukan aturan asosiasi antara suatu kombinasi item dalam suatu waktu.
2. *Sequence*, proses untuk menemukan aturan asosiasi suatu kobinasi item dalam suatu waktu dan diterapkan lebih dari satu periode.
3. *Clustering*, adalah proses pengelompokan sejumlah data/objek kedalam kelompok data sehingga setiap kelompok berisi data yang mirip.
4. *Classification*, proses penemuan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan dari suatu objek yang labelnya tidak diketahui.
5. *Regression*, adalah proses pemetaan data dalam suatu nilai prediksi.
6. *Forecasting*, adalah proses pengestimasiian nilai prediksi berdasarkan pola-pola didalam sekumpulan data.
7. *Solution*, adalah proses penemuan akar masalah dan *problem solving* dari persoalan bisnis yang dihadapi atau paling tidak sebagai informasi dalam pengambilan keputusan.

#### 2.2.4 Kategori *Data Mining*

Menurut (Vulandari, 2017), *data mining* dibagi menjadi dua kategori utama (hand dan Kamber, 2006) yaitu:

1) Prediktif

Tujuan dari tugas prediktif adalah untuk memprediksi nilai dari atribut tertentu berdasarkan pada nilai atribut lain. Atribut yang diprediksi biasanya dikenal sebagai target atau variabel tak bebas, sedangkan atribut-atribut yang digunakan untuk membuat prediksi dikenal sebagai *explanatory* atau variabel bebas.

2) Deskriptif

Tujuan dari tugas deskriptif adalah untuk menurunkan pola-pola (korelasi, *trend*, *cluster*, teori, dan *anomaly*) yang meringkas hubungan kecocokan dalam suatu data. Tugas *data mining* deskriptif merupakan penyelidikan dan sering kali memerlukan teknik *post-processing* untuk validasi dan penjelasan hasil.

#### 2.2.5 Manfaat *Data Mining*

Menurut (Vulandari, 2017), Pemanfaatan *data mining* dilihat dari dua sudut pandang, yaitu sudut pandang komersial dan sudut pandang keilmuan.

- a) Sudut pandang komersial, pemanfaatan *data mining* dapat digunakan untuk menangani meledaknya *volume* data, dengan menggunakan teknik komputasi dapat digunakan untuk menghasilkan informasi-informasi yang dibutuhkan yang merupakan *asset* yang dapat meningkatkan daya saing suatu institusi.

- b) Sudut pandang keilmuan, *data mining* dapat digunakan untuk *capture*, menganalisis serta menyimpan data yang bersifat *real time* dan sangat besar, misalnya :
1. *Remote sensor* yang ditempatkan pada suatu satelit.
  2. *Telescope* yang digunakan untuk memindai langit.
  3. Simulasi saintifik yang membangkitkan data dalam ukuran *terabytes*.

### 2.2.6 Penerapan *Data Mining*

Menurut (Vulandari, 2017), Penerapan *data mining* dibagi menjadi tiga bagian yaitu sebagai berikut :

1. Analisa pasar dan kumpulan sumber data yang digunakan seperti transaksi kartu kredit, kartu anggota *club* tertentu, kupon diskon, keluhan pembeli, ditambah dengan *study* dengan gaya hidup *public*. Beberapa solusi yang dapat diselesaikan dengan *data mining* antara lain (menebak target pasar, melihat pola beli pemakai dari waktu ke waktu, *cross market* analisis, *profil customer*, identifikasi kebutuhan *customer*, menilai loyalitas *customer*, informasi *summary*)
2. Analisa perusahaan dan *management* resiko
  - a) Perencanaan keuangan dan evaluasi *asset*, *data mining* dapat membantu melakukan, analisis dan prediksi *cash flow* serta dapat melakukan *contigent claim analysis* untuk mengevaluasi *asset*.
  - b) Perencanaan sumber daya, dengan melihat ringkasan informasi serta pola pembelanjaan dan pemasukan dari masing-masing *resource*.

- c) Persaingan, *data mining* dapat membantu untuk memonitor pesaing-pesaing dengan melihat *market direction* mereka.
3. Telekomunikasi, *data mining* melihat jutaan transaksi yang masuk dan melihat transaksi mana saja kah yang masih harus ditangani secara manual.

### **2.3 Metode Data Mining**

Menurut (Robi yanto, 2015) Analisis asosiasi atau *association rule mining* adalah teknik *data mining* untuk menemukan aturan asosiasi antara suatu kombinasi *item interestingness measure* yang dapat digunakan dalam *data mining* adalah:

1. *Support*, adalah suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu *item* atau *itemset* dari keseluruhan transaksi.
2. *Confidence*, adalah suatu ukuran yang menunjukkan hubungan antar dua *item* secara *conditional* (berdasarkan suatu kondisi tertentu).

#### **2.3.1 Algoritma Apriori**

Menurut (Robi yanto, 2015) *Algoritma apriori* adalah *algoritma* pengambilan data dengan aturan asosiatif (*Association rule*) untuk menentukan hubungan asosiatif suatu kombinasi item.

Menurut (Irsyad Djamiludin, 2017) *Algoritma Apriori* adalah salah satu *algoritma* yang melakukan pencarian *frequent itemset* dengan menggunakan teknik *association rule*. *Algoritma Apriori* menggunakan pengetahuan sebelumnya dari suatu *itemset* dengan frekuensi kemunculan yang sering atau juga yang bisa disebut frekuensi *itemset*.

### 2.3.2 Analisis Pola Frekuensi Tinggi dengan *Algoritma Apriori*

Mencari suatu nilai frekuensi *item* yang memenuhi syarat minimum dari nilai *support* dalam basis data. Nilai *support* sebuah *item* diperoleh dengan menggunakan rumus berikut:

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

**Rumus 2. 1** Frekuensi tinggi 1 *Itemset*

Nilai *support* dari kombinasi 2 *item* diperoleh dengan menggunakan rumus:

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

$$\text{Support (A,B)} = \frac{\Sigma \text{transaksi mengandung A dan B}}{\Sigma \text{Transaksi}} \times 100\%$$

**Rumus 2. 2** Menghitung kombinasi 2 *Itemset*

### 2.3.3 Pembentukan Aturan Asosiasi

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif A U B. Nilai *confidence* dari aturan A U B diperoleh dengan rumus berikut.

$$\text{Confidence} = P(B|A) = \frac{\Sigma \text{transaksi mengandung A dan B}}{\Sigma \text{Transaksi mengandung A}} \times 100\%$$

**Rumus 2. 3** Menghitung nilai *Confidence*

Untuk menentukan aturan asosiasi yang akan dipilih maka harus diurutkan berdasarkan *Support* dan *Confidence*. Aturan diambil sebanyak jumlah aturan yang memiliki hasil terbesar.

## 2.4 *Software Pendukung*

### 2.4.1 *Software Tanagra*

*Tanagra* adalah perangkat lunak pengolahan data, sumber terbuka, ramah penggunaan dan mudah untuk pembacaan hasil keputusan yang dibuat bagi siswa dan peneliti hingga pembisnis untuk menambang jumlah data yang sangat banyak bagi mereka. *Software* ini menyediakan beberapa metode penambangan dari data analisis, data eksplorasi, pembelajaran statistik, pembelajaran mesin dan bidang basis data. *Software* ini diprogram dalam Bahasa *pascal*. *Software* ini memiliki berbagai komponen seperti visualisasi data, *statistics*, *clustering*, *association*, *analysis factorial*, *regression* dan banyak lagi. Dengan menggunakan *Tanagra*, peneliti dapat memvisualisasikan suatu data menjadi lebih baik. Visualisasi data yang ada disajikan dalam bentuk *dataset*, memplot nilai pada grafik dan sebar plot. Ini digunakan untuk menunjukkan hubungan antara atribut dalam sumbu 2D. *Tanagra* mencakup *algoritma* pengelompokan dasar seperti *K-Means*, *EM – clustering* dll. Dalam basis data (Jain Vishal, 2013).



**Gambar 2. 2** *Tanagra Software*

**Sumber:** *Tanagra software logo*

### 2.4.2 *Microsoft Excel*

*Aplikasi Microsoft excel 2013* adalah aplikasi pengolahan data yang menyajikan informasi dengan baik dan digunakan oleh berbagai kalangan

contohnya, dosen, mahasiswa, siswa, serta perusahaan-perusahaan besar untuk mempresentasikan hasil produknya. Dengan *Microsoft excel 2013* penyampaian informasi apapun akan mudah untuk dipahami. (Auliya Rahman, Fitriah Yuridka, 2015).

## 2.5 Penelitian Terdahulu

Penelitian terdahulu merupakan acuan dalam sebuah penelitian, ada beberapa penelitian terdahulu yang digunakan dalam penelitian ini sebagai berikut.

1. Menurut Alfannisa Annurullah fajrin dan Algifanri Maulana tahun 2018 yang berjudul **“PENERAPAN DATA MINING UNTUK ANALISIS POLA PEMBELIAN KONSUMEN DENGAN ALGORITMA FP-GROWTH PADA DATA TRANSAKSI PENJUALAN SPARE PART MOTOR”** Pada perusahaan yang memiliki banyak cabang atau dealer seperti CV.TJAJAJA BARU ini, membutuhkan lokasi yang harus diperhatikan dalam mendirikan sebuah cabang baru, harus dapat mempengaruhi pola pembelian konsumen, karena pola pembelian setiap konsumen berbeda-beda. Hal ini perlu dianalisis lebih mendalam sehingga dapat menghasilkan informasi yang bermanfaat, serta memaksimalkan keuntungan yang bisa diperoleh. *Data Mining* dapat digunakan oleh perusahaan besar untuk menggali data serta menemukan informasi yang dapat menunjang dan meningkatkan suatu proses keberlangsungan bisnis perusahaan tersebut. Maka dalam penelitian ini dilakukan pengujian menggunakan *algoritma FP-*

*Growth* untuk membantu perusahaan menemukan pola pembelian konsumen dan transaksi penjualan *spare part*.

2. Menurut Irsyad Djamaludin dan Agus Nursikuwagus tahun 2017 yang berjudul **“ANALISIS POLA PEMBELIAN KONSUMEN PADA TRANSAKSI PENJUALAN MENGGUNAKAN ALGORITMA APRIORI”** Penambahan data perlu dilakukan jika informasi disimpan dalam suatu penyimpanan. Aktivitas ini menjadi dasar utama untuk menentukan suatu keputusan. Aktivitas yang sudah berjalan menemukan informasi transaksi penjualan dan pembelian yang hanya tersimpan pada *repository* sebagai suatu laporan saja. Sehingga informasi hanya menjadi data histori yang tidak digunakan. Berdasarkan kejadian ini, maka diperlukan suatu analisis untuk mendukung suatu keputusan. Analisis ini adalah teknik *mining* dengan menggunakan *algoritma Apriori*. Tujuan dari penelitian ini adalah mendapatkan pola penjualan dan pembelian suatu produk. Untuk menjawab tujuan maka digunakan *algoritma Apriori* dengan aturan asosiasi. Penentuan *itemset* yang digunakan adalah berjumlah 1, 2, dan 3 *itemset* produk. Sedangkan untuk *Minimum Support* yang digunakan adalah 30% dari jumlah transaksi yang ada. Untuk asosiasi antar produk ditentukan dengan *confidence* sebesar 70%. Penelitian ini diselesaikan dengan menggunakan *UML* yang terdiri dari *Usecase Diagram*, *Class Diagram*, dan *Activity Diagram*. Hasil dari penelitian ini adalah perangkat lunak analisis pola penjualan dan pembelian konsumen dengan *Algoritma Apriori*. Sedangkan keputusan yang diperoleh adalah asosiasi antar produk dengan menggunakan nilai *confidence*.

Untuk *Final Rule Association* adalah produk bolu *cake meses* dan bolu *cake keju* dengan nilai *confidence* sebesar 84,62%.

3. Menurut Dewi Kartika Pane tahun 2013 yang berjudul **“IMPLEMENTASI DATA MINING PADA PENJUALAN PRODUK ELEKTRONIK DENGAN ALGORITMA APRIORI (STUDI KASUS : KREDITPLUS)”**  
 Penjualan produk elektronik, khususnya laptop mengalami peningkatan setiap bulannya, produk yang ditawarkan bermacam merek, merek mempengaruhi masyarakat untuk membeli produk tersebut, untuk mengetahui merek dengan penjualan terbanyak diperlukan *algoritma apriori* untuk dapat mengetahuinya, dan dengan bantuan tools *Weka*, produk dengan penjualan terbanyak dapat diketahui. *Algoritma apriori* termasuk jenis aturan asosiasi pada *data mining*. Salah satu tahap analisis asosiasi yang menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien adalah analisis pola frekuensi tinggi (*frequent pattern mining*). Penting tidaknya suatu asosiasi dapat diketahui dengan dua tolak ukur, yaitu : *support* dan *confidence*. *Support* (nilai penunjang) adalah persentase kombinasi *item* tersebut dalam database, sedangkan *confidence* (nilai kepastian) adalah kuatnya hubungan antar-*item* dalam aturan asosiasi. *Algoritma apriori* dapat membantu untuk pengembangan strategi pemasaran.
4. Menurut Robi Yanto dan Riri Khoiriah tahun 2015 yang berjudul **“IMPLEMENTASI DATA MINING DENGAN METODE ALGORITMA APRIORI DALAM MENENTUKAN POLA PEMBELIAN OBAT”** *Data mining* merupakan proses untuk mendapatkan informasi yang berguna dari

gudang basis data yang berupa ilmu pengetahuan. penelitian ini melakukan analisa data dengan menggunakan *data mining* dan metode *algoritma apriori*. Sistem yang dibangun ditujukan untuk pemenuhan dalam penentuan pola pembelian obat dengan menggunakan bahasa pemrograman *Visual Basic 6.0* dan *database Mysql* pada studi kasus di sektor kesehatan. Sistem ini dibangun berdasarkan kebutuhan pengguna yang diperoleh melalui metode wawancara dan studi lapangan. Metodologi pengembangan sistem yang digunakan yaitu metode *waterfall* yang terdiri analisis, desain, pengkodean dan pengujian. Hasil pengujian dengan *algoritma apriori* dan sistem yang dibangun menunjukkan hasil yang telah memenuhi kebutuhan dalam penentuan pola pembelian obat berdasarkan kecenderungan pembelian obat oleh pelanggan. Dibandingkan dengan sistem yang sedang berjalan kinerja tersebut ditunjukkan pada efektifitas informasi dari sistem tentang penentuan pola pembelian obat untuk ketersediaan obat dan tata letak obat untuk memudahkan dalam mengetahui keberadaan obat yang dilihat dari 2 *itemset* obat.

5. Menurut Goldie Gunadi Dana Indra Sensuse tahun 2012 yang berjudul **“PENERAPAN METODE DATA MINING MARKET BASKET ANALYSIS TERHADAP DATA PENJUALAN PRODUK BUKU DENGAN MENGGUNAKAN ALGORITMA APRIORI DAN FREQUENT PATTERN GROWTH (FP-GROWTH) : STUDI KASUS PERCETAKAN PT. GRAMEDIA”** *Data mining* merupakan proses analisa data untuk menemukan suatu pola dari kumpulan data tersebut. *Data mining*

mampu menganalisa data yang besar menjadi informasi berupa pola yang mempunyai arti bagi pendukung keputusan. Salah satu teknik *data mining* yang dapat digunakan adalah *association data mining* atau yang biasa disebut dengan istilah *market basket analysis*. *Market basket* didefinisikan sebagai suatu *itemset* yang dibeli secara bersamaan oleh pelanggan dalam suatu transaksi. *Market basket analysis* adalah suatu alat yang ampuh untuk pelaksanaan strategi *cross-selling*. Metode ini dimulai dengan mencari sejumlah *frequent itemset* dan dilanjutkan dengan pembentukan aturan-aturan asosiasi (*association rules*). *Algoritma Apriori* dan *frequent pattern growth (FP-growth)* adalah dua algoritma yang sangat populer untuk menemukan sejumlah *frequent itemset* dari data-data transaksi yang tersimpan dalam basis data. Dalam penelitian ini *algoritma Apriori* dan *frequent pattern growth (FP-growth)* digunakan untuk membantu menemukan sejumlah aturan asosiasi dari basis data transaksi penjualan produk buku di Percetakan PT. Gramedia, sehingga untuk selanjutnya dapat digunakan sebagai pertimbangan dalam membuat strategi pemasaran dan penjualan yang efektif.

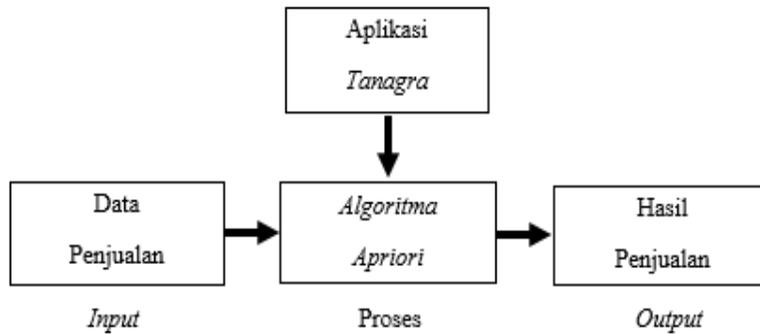
6. Menurut Jiao Yabing tahun 2013 yang berjudul ***“Research of an Improved Apriori Algorithm in Data Mining Association Rules”*** *Algoritma Apriori* adalah algoritma dari aturan asosiasi, yang menyebutkan semua set *item* yang sering muncul. Saat ini *algoritma apriori* mengalami data yang padat karena besar sejumlah pola panjang muncul, kinerja *algoritma* ini menurun drastis. Untuk menemukan aturan yang lebih baik, penulis mengusulkan adanya peningkatan *algoritma* aturan asosiasi, kepada *algoritma Apriori* klasik yang

pada akhirnya, algoritma yang ditingkatkan akan diverifikasi, kemudian hasilnya akan menunjukkan bahwa peningkatan *algoritma apriori* lebih efektif, dapat mengekstraksi lebih banyak informasi nilai maupun data.

7. Menurut Wiwit Agus Triyanto tahun 2014 yang berjudul “**ASSOCIATION RULE MINING UNTUK PENENTUAN REKOMENDASI PROMOSI PRODUK**” Rekomendasi Promosi Produk merupakan model aplikasi dari hasil observasi terhadap keadaan dan keinginan pelanggan untuk membeli suatu produk. *Association Rule Mining* dapat memberikan dukungan keputusan antara barang yang dibeli secara bersama-sama oleh pelanggan. Penemuan pola barang yang dibeli oleh pelanggan sangat penting dikarenakan dapat membantu rekomendasi promosi produk sehingga strategi pemasaran menjadi lebih tepat sasaran. Penelitian ini menggunakan *Algoritma FP-Growth* untuk pendekatan asosiasi. Hasil dari penelitian ini adalah *rule* yang dapat digunakan untuk membantu menentukan rekomendasi promosi suatu produk secara lebih tepat, yaitu dengan menghasilkan 3 *rule* ketika menggunakan *minimum support* 40% dan *minimum confidence* 80%.

## 2.6 Kerangka Pemikiran

Konsep yang dikembangkan dalam penelitian ini sebagai berikut :



**Gambar 2. 3** Kerangka Pemikiran

**Sumber:** Data Penelitian (2019)

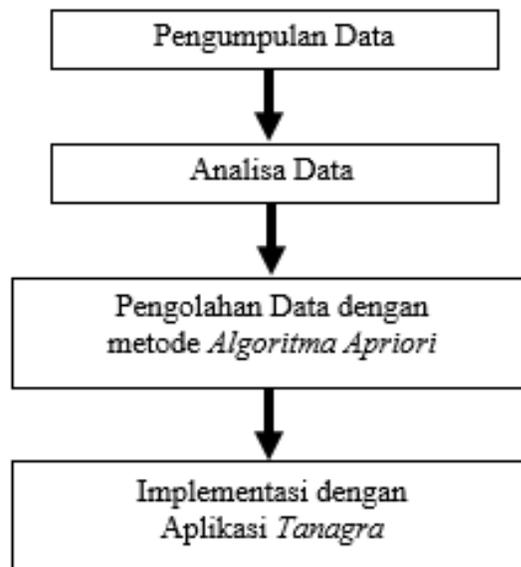
Data penjualan yang bersumber dari data transaksi penjualan rumah yang digunakan sebagai *input*, kemudian data dikelola dengan menggunakan metode *algoritma apriori* sehingga mendapat hasil atau *output* yang digunakan sebagai pengembangan dan peningkatan pembangunan rumah. Data penjualan yang bersumber data transaksi mencakup tiga variabel yaitu tipe rumah, luas tanah, harga jual, dan hasil akhir penjualan yang dikumpulkan dalam waktu tiga tahun terakhir oktober 2016 sampai dengan september 2019 dengan data yang diambil dari data perbulannya selama tiga puluh enam bulan. Algoritma apriori yang digunakan akan membantu kita menambang suatu data yang banyak sehingga menghasilkan aturan dan pengetahuan baru.

## BAB III

### METODE PENELITIAN

#### 3.1 Desain Penelitian

Dalam proses penelitian, terlebih dahulu perlu dibuat desain penelitian. Hal ini bertujuan untuk memberikan kemudahan penelitian yang lebih lanjut. Adapun desain dalam penelitian ini adalah sebagai berikut:



**Gambar 3. 1** Desain Penelitian

**Sumber:** Data Penelitian (2019)

Berdasarkan desain penelitian diatas yaitu suatu alur penelitian yang akan dilakukan dalam proses penelitian ini. Desain penelitian diatas menunjukkan proses dari awal kegiatan penelitian sampai akhir penelitian.

Berikut ini adalah penjelasan dari desain penelitian diatas:

1. Pengumpulan Data

Pengumpulan data dalam penelitian ini menggunakan teknik observasi yaitu observasi langsung ke lokasi penelitian, selanjutnya melakukan wawancara dengan pihak *accounting PT. Glory Point* dan terakhir melakukan studi pustaka mengenai materi-materi *data mining* dan metode *algoritma apriori* yang diperoleh dari buku-buku dan jurnal nasional maupun internasional.

2. Analisa Data

Analisa data yaitu proses dimana data yang diperoleh untuk menentukan operasional variabel yang dijadikan *input* dalam proses penelitian serta *output* yang dihasilkan sesuai dengan apa yang diinginkan.

3. Pengolahan Data Dengan Metode *Algoritma Apriori*

Data yang telah di analisa selanjutnya akan diolah menggunakan teknik *data mining* menggunakan metode *algoritma apriori*. Data yang diolah mencakup dua bagian yaitu tipe rumah dan total penjualan rumah yang laku. Untuk mendukung suatu keputusan pengolahan peneliti juga menambahkan luas tanah dan harga pada kolom data yang akan diteliti sebagai informasi tambahan dalam pengambilan keputusan, dan tidak termasuk bagian yang diolah.

4. Implementasi Dengan Aplikasi *Tanagra*

Setelah semua data di olah maka data tersebut diimplementasikan kedalam aplikasi *Tanagra*. Data yang diolah dengan aplikasi ini akan menjadi pertimbangan dalam perhitungan yang dilakukan secara manual.

## **3.2 Metode Pengumpulan Data**

Dari data yang akan di teliti, teknik pengumpulan data pada penelitian ini menggunakan tiga teknik yaitu teknik observasi, wawancara dan studi pustaka. Metode ini dilakukan secara berkala selama proses penelitian.

### **3.2.1 Observasi**

Merupakan serangkaian proses biologis dan psikologis yang tersusun secara kompleks. Dua diantara terpenting adalah proses-proses pengamatan dan ingatan. Observasi dalam penelitian ini dilakukan secara langsung di *PT. Glory Point* untuk melakukan pengamatan terhadap data penjualan rumah.

### **3.2.2 Wawancara**

Merupakan salah satu instrumen yang digunakan untuk menggali data secara lisan. Hal ini haruslah dilakukan secara mendalam agar mendapatkan data yang *valid* dan detail. Wawancara dilakukan langsung pada pihak *accounting PT. Glory Point*.

### **3.2.3 Studi Pustaka**

Dalam penelitian ini peneliti melakukan studi pustaka yang berhubungan dengan tema penelitian & metode penelitian yaitu jurnal dan buku yang berhubungan, untuk mencari informasi menyusun teori-teori yang berhubungan dengan pembahasan.

### 3.3 Operasional Variabel

Variabel penelitian adalah suatu atribut atau sifat atau nilai dari orang, objek, organisasi atau kegiatan yang mempunyai variasi tertentu yang ditetapkan oleh peneliti untuk dipelajari dan kemudian ditarik kesimpulannya.

Operasional variabel dalam penelitian ini dapat dijelaskan dan diuraikan sebagai berikut:

#### 1. Tipe Rumah

Tipe rumah merupakan tipe-tipe rumah yang paling banyak diminati dan dibeli oleh konsumen dan merupakan tipe rumah sebagai pemilihan stok. Tipe rumah pada PT. Glory Point adalah Tipe 36/72a, Tipe 60/59, Tipe 46/84, Tipe 84 yang terletak di lokasi tanjung riau dan Tipe 36/72b, Tipe 42/78, Tipe 52/91, Tipe 36/96, Tipe 52/105, Tipe 105 yang terletak dilokasi Tiban Koperasi.

#### 2. Harga Rumah

Disediakan untuk menunjang perbandingan harga yang ada dalam penjualan rumah terhadap harga rumah lainnya yang berbeda tipe. Sehingga dapat menjadi pembanding pada tipe lainnya. Harga yang variasi menyesuaikan lokasi dan luas tanah yang ada.

#### 3. Luas tanah

Disediakan untuk menunjang perbandingan tipe rumah dan harga rumah dalam transaksi pemilihan rumah yang mana luas tanah akan dijadikan salah satu kesimpulan bahwa luas tanah mempengaruhi jual beli rumah dan minat konsumen didalam memilihnya. Terdapat Tipe rumah 36/72a dengan luas

tanah 72m<sup>2</sup>, Tipe rumah 36/72b dengan luas tanah 72m<sup>2</sup>, Tipe rumah 36/96 dengan luas tanah 98,32m<sup>2</sup>, Tipe rumah 60/59 dengan luas tanah 95.8m<sup>2</sup>, Tipe rumah 42/78 dengan luas tanah 78m<sup>2</sup>, Tipe rumah 46/84 dengan luas tanah 84m<sup>2</sup>, Tipe rumah 52/91 dengan luas tanah 101m<sup>2</sup>, Tipe rumah 52/105 dengan luas tanah 105m<sup>2</sup>, Tipe rumah 84 dengan luas tanah 85,54m<sup>2</sup>, Tipe rumah 105 dengan luas tanah 110m<sup>2</sup>.

**Tabel 3. 1** Operasional Variabel

<b>No.</b>	<b>Tipe Rumah</b>	<b>Luas Tanah (m2)</b>	<b>Harga</b>
1	36/72a	72	Rp 230,000,000
2	36/72b	72	Rp 400,000,000
3	36/96	98.32	Rp 301,170,000
4	42/78	78	Rp 480,000,000
5	46/84	84	Rp 330,400,000
6	52/91	101	Rp 520,000,000
7	52/105	105	Rp 675,000,000
8	60/59	95.8	Rp 416,800,000
9	84	85.54	Rp 667,555,000
10	105	110	Rp 782,500,000

**Sumber:** Data Penelitian (2019)

### 3.4 Metode Perancangan Sistem

Dalam penelitian ini menggunakan metode *algoritma apriori*. Langkah-langkah perancangan sistem dalam penelitian ini adalah sebagai berikut:



**Gambar 3. 2** Metode Perancangan Sistem

**Sumber:** Data Penelitian (2019)

Dari gambar 3.2 diatas dapat dijelaskan proses awal hingga proses akhir. Dimulai dari data penjualan dari *PT. Glory Point* yang telah dianalisa oleh peneliti, data yang diperoleh yaitu data penjualan rumah di *PT. Glory Point* dari Oktober 2016 sampai September 2019. Selanjutnya data yang tersebut dimasukkan kedalam tabel tabular, kemudian data-data akan dicari nilai *minimum support* dari 1 *itemset* hingga akhir yaitu 5 *itemset*, setelah nilai *minimum support* diperoleh maka langkah selanjutnya yaitu menentukan nilai *confidence* sampai menghasilkan nilai *final rule*nya yang dilakukan secara manual menggunakan metode *algoritma apriori*. Kemudian tahap terakhir yaitu melakukan pengujian dengan menggunakan

aplikasi *tanagra*, sebagai pembuktian bahwa hasil yang diperoleh secara manual sama dengan hasil implementasi dengan aplikasi.

### 3.5 Lokasi Dan Jadwal Penelitian

#### 3.5.1 Lokasi Penelitian

Penelitian ini dilakukan di *PT. Glory Point* yang beralamat di Komp. Ruko *Orchid Business Center* Blok C1 No. 5 Batam Center, Batam.

#### 3.5.2 Jadwal Penelitian

Penelitian ini dilakukan dalam waktu satu semester seperti yang telah dijadwalkan didalam kalender akademik, setelah pengajuan judul penelitian ini diterima. Penelitian ini dimulai dari bulan september 2019 hingga bulan februari 2020. Jadwal penelitian dalam penelitian ini dapat dilihat ditabel berikut:

**Tabel 3. 2** Jadwal Penelitian

Kegiatan	Waktu Kegiatan																						
	Sept 2019			Okt 2019				Nov 2019				Des 2019				Jan 2020				Feb 2020			
	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Pengajuan Judul	■	■	■																				
Penyusunan BAB I				■	■	■	■																
Penyusunan BAB II								■	■	■	■												
Penyusunan BAB III												■	■	■	■								
Penyusunan BAB IV																■	■	■	■				
Penyusunan BAB V																				■	■		
Pengumpulan Skripsi																						■	■

**Sumber:** Data Penelitian 2019