

**DATA MINING SISTEM TATA LETAK MATERIAL DI
PT BATAM CYCLECT**

SKRIPSI



Oleh
Yuni Shantika Hutapea
160210181

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2021**

**DATA MINING SISTEM TATA LETAK MATERIAL DI
PT BATAM CYCLECT**

SKRIPSI



**Oleh
Yuni Shantika Hutapea
160210181**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2021**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Yuni Shantika Hutapea
NPM : 160210181
Fakultas : Teknik dan Komputer
Program Studi : Teknik Informatika

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

“DATA MINING SISTEM TATA LETAK MATERIAL DI PT BATAM CYCLECT”

Adalah hasil karya sendiri dan bukan “Duplikasi” darikarya orang lain. Sepengetahuan saya, didalam naskah “Data Mining Sistem Tata Letak Material Di PT Batam Cyclect” ini tidak terdapat kerja 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 terdapat didalam naskah “Data Mining Sistem Tata Letak Material Di PT Batam Cyclect” ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah “Data Mining Sistem Tata Letak Material Di PT Batam Cyclect” ini digugurkan dan sarjana computer 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.



***DATA MINING* SISTEM TATA LETAK MATERIAL DI
PT BATAM CYCLECT**

SKRIPSI

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

**Oleh
Yuni Shantika Hutapea
160210181**

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

Batam, 28 Januari 2021

A handwritten signature in black ink, appearing to be 'Rahmat Fauzi', written over a horizontal line.

Rahmat Fauzi, S.Kom., M.Kom.

Pembimbing

ABSTRAK

PT BATAM CYCLECT atau lebih dikenal dengan Cyclect adalah salah satu perusahaan yang bergerak dalam bidang elektrik di *shipyard*. PT BATAM CYCLECT digudang materialnya kurang baik prosedur penataan material pada gudang yang mengakibatkan gudang terkesan sempit dan kurang teratur. Maka dibuat sebuah sistem yang membantu dalam mengolah data tata letak material berdasarkan tipe barang seperti, material *consumable*, material *electric*, material *welding*, *tools electric*, material kimia dan banyak jenis lainnya. *Data Mining* merupakan solusi dalam mengembangkan usaha atau cara memanfaatkan data untuk proses prediksi kedepannya. Metode yang digunakan adalah *association Rules* untuk menemukan aturan assosiatif yang saling terhubung dengan item satu dan lainnya. Algoritma Apriori adalah algoritma yang digunakan untuk menghasilkan *association Rules* dan terkenal dalam menentukan pola *frequency* tinggi. Tujuannya adalah untuk mempermudah *storeman* dalam pelayan dan mempersingkat waktu pada saat pencarian barang secara bersamaan dengan sistem tata letak material. Algoritma apriopri memanfaatkan proses tersebut untuk mengurangi atau mempersempit ruang pencarian frequentsi kandidat *itemset*. Untuk menguji akurasi algoritma apriori menggunakan software RapidMiner. Hasil dari penelitian yang dilakukan maka diperoleh 10 rule asosiasi dengan nilai minimal support 5% dan confidence 50% dengan hasil 3-itemset yaitu jika memesan Marking Tape dan marker maka akan memesan PVC Insulation Tape dengan nilai Support 15% dan nilai confidence 89%.

Kata Kunci: Algoritma Apriori, Association Rule, Data Mining, RapidMiner, Tata Letak

ABSTRACT

PT BATAM CYCLECT or better known as Cyclect is a company engaged in the electrical sector in the shipyard. PT BATAM CYCLECT in its material warehouse lacks good material arrangement procedures in the warehouse which results in the warehouse looking cramped and less orderly. So a system is created that helps in processing material layout data based on the type of item, such as material consumables, electric materials, welding materials, electric tools, chemical materials and many other types. Data Mining is a solution in developing a business or a way to use data for future prediction processes. The method used is association rules to find associative rules that are connected with one item and another. The Apriori Algorithm is an algorithm used to generate association rules and is well-known in determining high frequency patterns. The goal is to make it easier for the storeman in the waiter and shorten the time when searching for goods simultaneously with the material layout system. The Apriori algorithm takes advantage of this process to reduce or narrow the search space for itemset candidate frequency. To test the accuracy of the Apriori algorithm using the Rapid Miner software. The results of the research carried out obtained 10 association rules with a minimum support value of 5% and 50% confidence with 3-itemset results, namely if you order Marking Tape and markers you will order PVC Insulation Tape with a Support value of 15% and a confidence value of 89%.

Keywords: *Apriori Algorithm, Association Rule, Data Mining, Rapid Miner, Layout*

KATA PENGANTAR

Puji syukur kehadiran Tuhan Yang Maha Esa yang telah melimpahkan segala rahmat dan kurnia-Nya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu syarat untuk menyelesaikan program studi strata satu (S1) pada program Studi Teknik Informatika Universitas Putera Batam.

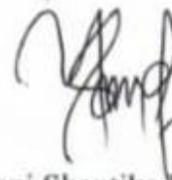
Peneliti menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati. Dengan segala keterbatasan, penulis menyadari pula bahwa skripsi ini tidak akan selesai tanpa bantuan, bimbingan dan dorongan dari berbagai pihak. Untuk itu dengan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

1. Ibu Dr. Nur Elfi Husda, S.Kom.,M.SI. Selaku Rektor Universitas Putera Batam.
2. Ketua Program Studi Teknik Informatika, Bapak Andi Maslan, S.T., M.SI.
3. Kepada Bapak Rahmat Fauzi, S.Kom., M.Kom. Selaku pembimbing skripsi pada program Studi Teknik Informatika Universitas Putera Batam.
4. Dosen dan Staff Universitas Putera Batam.
5. Kedua orang tua, abang dan adik yang selalu memberikan dukungan dan nasehat yang penuh kepada penulis.
6. Seluruh keluarga besar yang selalu member nasehat-nasehat dan motivasi.
7. PT Batam Cyclelect yang telah memberikan izin pengumpulan data penelitian.

8. Kepada Alvan Christnandio, rekan kerja saya yang tidak pernah berhenti memberikan motivasi serta dukungan selama pengerjaan skripsi ini.
9. Kepada Pye Sone Phyo, Boss saya yang selalu memberi masukan dan memberi bimbingan untuk menyelesaikan skripsi ini.
10. Para sahabat seperjuangan yang selalu setia mendampingi dan memberikan dukungan serta semangat. Mereka adalah Rizki Kurniati, Sri Wahyuni, Khoirul Anam, Rizal Arisandy, Andriani Pratiwi, Joshua Gumelar S, Doli Candra, Theresia E.B.

Semoga Tuhan Yang Maha Esa membalas kebaikan dan selalumencurahkan rahmat dan karunia-Nya.Amin.

Batam, 28 Januari 2021



Yuni Shantika Hutapea

Penulis

DAFTAR ISI

	Halaman
HALAMAN SAMPUL DEPAN	i
HALAMAN JUDUL	ii
SURAT PERNYATAAN ORISINALITAS	iii
HALAMAN PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	5
1.3 Pembatasan Masalah	5
1.4 Perumusan Masalah.....	6
1.5 Tujuan Penelitian.....	6
1.6 Manfaat Penelitian.....	7
1.6.1 Manfaat Teoritis.....	7
1.6.2 Manfaat Praktis	7
BAB II	9
KAJIAN PUSTAKA	9
2.1 <i>Knowledge Discovery In Database (KDD)</i>	9

2.1.1	Proses <i>Knowledge Discovery In Database</i>	10
2.2	Data Mining.....	12
2.2.1	Sejarah <i>Data Mining</i>	12
2.2.2	Pengertian Data Mining	14
2.2.3	Manfaat Data Mining.....	15
2.2.4	Strategi <i>Data Mining</i>	16
2.3	Metode <i>Data Mining</i>	17
2.4	Algoritma <i>Data Mining</i>	19
2.4.1	<i>Association Rules</i>	19
2.4.2	Algoritma <i>Apriori</i>	22
2.4.3	Algoritma C4.5	26
2.4.4	Algoritma <i>FP-Growth</i>	26
2.5	<i>Market Basket Analysis</i>	27
2.6	<i>Software</i> Pendukung.....	28
2.6.1	<i>RapidMiner</i>	28
2.7	Tata Letak Barang	29
2.8	Penelitian Terdahulu.....	29
2.9	Kerangka Pemikiran	34
BAB III	36
METODE PENELITIAN	36
3.1	Desain Penelitian	36
3.2	Teknik Pengumpulan Data	38

3.3	Operasional Variable	39
3.4	Metode Analisis dan Rancangan Sistem	41
3.5	Lokasi dan Jadwal Penelitian	43
3.5.1	Lokasi Penelitian.....	43
3.5.2	Jadwal Penelitian	43
BAB IV		45
HASIL DAN PEMBAHASAN		45
4.1	Analisa Data	45
4.1.1	Pengumpulan data.....	45
4.1.2	<i>Cleansing Data</i>	46
4.1.3	Klasifikasi Data.....	48
4.1.4	<i>Transformation</i>	49
4.1.5	Perhitungan Algoritma Apriori.....	52
4.1.6	Pembentukan Itemset.....	53
4.1.7	Pembentukan Aturan Asosiasi	68
4.2	Hasil Pengujian.....	73
4.3	Visualisasi Tata Letak Barang.....	82
BAB V.....		85
SIMPULAN DAN SARAN		85
5.1	Kesimpulan.....	85
5.2	Saran.....	86
DAFTAR PUSTAKA		87

DAFTAR GAMBAR

	Halaman
Gambar 2. 1 Akar Bidang Ilmu.....	9
Gambar 2. 2 Proses <i>Knowledge Discovery in Database</i> (KDD)	11
Gambar 2. 3 Strategi Data Mining	16
Gambar 2. 4 Flowchart Algoritma Apriori	25
Gambar 2. 5 Logo RapidMiner	28
Gambar 2. 6 Tahapan Kerangka Pemikiran	34
Gambar 3. 1 Desain Penelitian.....	36
Gambar 3. 2 Metode Analisis Rancangan Sistem.....	41
Gambar 3. 3 Lokasi Penelitian	43

DAFTAR TABEL

	Halaman
Tabel 3. 1 Jadwal Pelaksanaan Penelitian.....	44
Tabel 4. 1 Data Mentah Penjualan	45
Tabel 4. 2 Rekapitulasi <i>Cleansing</i> Data.....	46
Tabel 4. 3 (Lanjutan) Rekapitulasi <i>Cleansing</i> Data.....	47
Tabel 4. 4 Klasifikasi Data.....	48
Tabel 4. 5 (Lanjutan) Klasifikasi Data.....	49
Tabel 4. 6 Tabulasi Data	50
Tabel 4. 7 Frekuensi Item	53
Tabel 4. 8 (Lanjutan) Frekuensi Item.....	53
Tabel 4. 9 Pola Frekuensi Tinggi 1- <i>Itemset</i>	57
Tabel 4. 10 Pembentukan 2- <i>Itemset</i>	58
Tabel 4. 11 Lanjutan Pembentukan 2- <i>Itemset</i>	59
Tabel 4. 12 Lanjutan Pembentukan 2- <i>Itemset</i>	60
Tabel 4. 13 Lanjutan Pembentukan 2- <i>Itemset</i>	61
Tabel 4. 14 Pola Frekuensi Tinggi 2- <i>Itemset</i>	62
Tabel 4. 15 Lanjutan Pola Frekuensi Tinggi 2- <i>Itemset</i>	63
Tabel 4. 16 Perhitungan <i>Support</i>	68
Tabel 4. 17 Perhitungan <i>Confidence</i>	69
Tabel 4. 18 Lanjutan Perhitungan <i>Confidence</i>	70
Tabel 4. 19 Lanjutan Perhitungan <i>Confidence</i>	71
Tabel 4. 20 Minimum <i>Confidence</i>	72

DAFTAR RUMUS

	Halaman
Rumus 2. 1 <i>Support 1-Itemset</i>	22
Rumus 2. 2 <i>Support 2-Itemset</i>	22
Rumus 2. 3 <i>Confidence</i>	22
Rumus 2. 4 <i>Lift Ratio</i>	24

BAB I

PENDAHULUAN

1.1 Latar Belakang

Gudang atau *storage* merupakan tempat menyimpan barang, baik bahan baku yang dapat diproses, ataupun barang jadi yang siap untuk dipasarkan. Gudang dapat menjadi sebuah sistem logistik dari sebuah perusahaan yang berfungsi untuk menyimpan produk dan menyediakan informasi mengenai status barang, kondisi barang atau produk yang disimpan sampai produk tersebut diminta sesuai dengan jadwal. *Storage* atau gudang adalah sebagian besar bahan baku yang harus bergerak untuk diproduksi jangan sampai barang yang ada di gudang disimpan atau tidak keluar terlalu lama karena akan mengakibatkan penambahan biaya penyimpanan dan tidak dapat menghasilkan pendapatan. Tujuan gudang adalah agar barang yang tersimpan tetap dalam kondisi baik saat akan di distribusikan kepada para pelanggan atau peminta stok pada waktu dan jumlah yang tepat.

PT BATAM CYCLECT atau lebih dikenal dengan Cyclelect adalah salah satu perusahaan yang bergerak dalam bidang elektrik di *shipyard*. Dalam sehari Cyclelect dapat melakukan *repair* atau perbaikan kapal lebih dari 5 kapal baik kapal pertamina, kapal barang, kapal tongkang dan masih banyak lagi. Setiap hari karyawan Cyclelect melakukan transaksi atau order material ke *store* yang menghasilkan data barang yang semakin bertambah setiap harinya, namun data tersebut tidak berguna apabila tidak diolah untuk mendapatkan data yang lebih mendalam. Tumpukan data tersebut dapat

menghasilkan suatu pola barang keluar dan barang masuk sehingga dari pola tersebut dapat diketahui barang *comsumble* atau kebutuhan pokok karyawan. Pola ini dapat digunakan untuk mengetahui sistem pencarian dan tata letak barang. Permasalahan yang dihadapi PT BATAM CYCLECT digudang material adalah kurang baiknya prosedur penataan material pada gudang yang mengakibatkan gudang terkesan sempit dan kurang teratur. Kondisi tata letak gudang yang tidak berdasarkan penataan barang menyebabkan operator gudang sulit dalam menangani material. Dalam proses persediaan (*stock*) dan permintaan material di Cyclelect sering terjadi keterlambatan dan kelebihan persediaan material karena tidak adanya inventori material bulanan yang mengakibatkan proses order material terlambat dan kelebihan persediaan material di gudang Cyclelect.

Untuk mengatasi permasalahan yang ada di PT BATAM CYCLECT akan dibuat sebuah sistem yang membantu dalam mengolah data tata letak material berdasarkan tipe barang seperti, material *consumable*, material *electric*, material *welding*, *tools electric*, material kimia dan banyak jenis lainnya. Metode yang cocok dengan permasalahan yang dihadapi oleh Cyclelect adalah metode *association Rule*. Mendeteksi kumpulan atribut yang muncul bersamaan (co-occur) dalam *frequency* yang sering dan membentuk sejumlah kaidah dari kumpulan-kumpulan tersebut (Hasibuan et al., 2017). Metode ini membantu operator gudang dalam menangani prosedur penataan material.

Data Mining merupakan solusi dalam mengembangkan usaha atau cara memanfaatkan data untuk proses prediksi kedepannya (Wulandari & Rahayu, 2014). *Data Mining* adalah serangkaian proses untuk menggali nilai tambah berupa informasi

yang selama ini tidak diketahui secara manual dari suatu *database* dengan melakukan penggalian pola-pola dari data dengan tujuan untuk memanipulasi data menjadi informasi yang lebih berharga yang diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat dalam *database*. *Data Mining* memiliki beberapa cara dalam pengolahan yaitu, *association Rule*, *classification*, *clustering*, *regression*, *forecasting* dan *sequence analysis* (Maulana & Fajrin, 2018).

Metode *association Rules* adalah suatu teknik data mining untuk menemukan hubungan kombinasi *item* barang. Contoh *association Rules* dari analisa operator gudang di *storage Cyclect* adalah dapat diketahui berapa besar kemungkinan seorang karyawan *Cyclect* untuk *order marking tape* bersamaan dengan *marker paint black plus yellow*. Dengan informasi ini *storeman* dapat mengatur penempatan material atau *layout* pada gudang tersebut, tujuannya adalah untuk mempermudah *storeman* dalam pelayan dan mempersingkat waktu pada saat pencarian barang secara bersamaan. Berdasarkan alasan ini peneliti tertarik untuk merancang sistem tata letak material. Penelitian ini bertujuan untuk menganalisa bagaimana proses penataan material dengan menggunakan data mining. Penelitian ini juga bermanfaat membantu operator gudang atau *storeman* mempermudah pencarian item material dan inventori persediaan bulanan maupun tahunan. Metode yang digunakan adalah *association Rules* untuk menemukan aturan asosiatif yang saling terhubung dengan item satu dan lainnya.

Algoritma Apriori adalah algoritma yang digunakan untuk menghasilkan *association Rules* dan terkenal dalam menentukan pola *frequency* tinggi. Apriori

menggunakan *k-itemset* untuk mengeksplorasi *itemset* dengan melakukan pendekatan iteraksi yang dikenal dengan pencarian *level-wise*. Algoritma apriopri memanfaatkan proses tersebut untuk mengurangi atau mempersempit ruang pencarian *frequency* kandidat *itemset*.

Dari hasil pengamatan penulis, gudang PT BATAM CYCLECT melakukan transaksi sangat tinggi, akan tetapi data transaksi tersebut hanya tersimpan sebagai laporan dan arsip oleh Perusahaan BATAM CYCLECT. Peletakan material di rak masih kurang tertata. Pengaturan tata letak material sangat mempengaruhi proses pencarian dari *storeman* dengan bantu tata letak yang strategi dan memiliki keterkaitan barang satu dengan yang lain. Untuk penyusunan tata letak material memiliki ketentuan dan criteria seperti barang yang *consumable*, material yang akan diorder secara bersamaan dan hubungan produk satu dengan produk lain.

Berdasarkan masalah diatas, maka diangkat sebuah penelitian yang berjudul “**DATA MINING SISTEM TATA LETAK MATERIAL DI PT BATAM CYCLECT** “.

1.2 Identifikasi Masalah

Berdasarkan latar belakang diatas, penulis mengidentifikasi beberapa masalah yang akan dijadikan bahan penelitian yaitu:

1. Sering terjadi kesalahan data stock material yang mengakibatkan persediaan material kurang atau habis.
2. *Storeman* sering kesulitan mendapatkan material yang dibutuhkan.
3. Sistem tata letak material PT BATAM CYCLECT masih tidak teratur dan terkesan sempit.
4. *Storeman* tidak melakukan inventori sehingga mengakibatkan terjadinya kekurangan atau kelebihan persediaan material didalam *storage*.
5. Banyaknya transaksi membuat data menumpuk dan tidak digunakan dengan maksimal.

1.3 Pembatasan Masalah

Agar pembahasan dalam penelitian ini tidak keluar dari tema penelitian dan judul penelitian, maka penulis memberikan batasan masalah adalah sebagai berikut :

1. Hanya fokus pada tata letak material di PT BATAM CYCLECT.
2. Data yang akan diolah adalah data transaksi pada bulan Oktober 2019 – September 2020.
3. Metode yang digunakan adalah metode asosiasi dengan menggunakan algoritma apriori.
4. Hasil dari penelitian ini akan dilakukan pengujian dengan aplikasi RapidMiner.

5. Variabel input yang digunakan berupa jenis item.

1.4 Perumusan Masalah

Berdasarkan latar belakang dan batasan masalah penelitian di atas, dapat dirumuskan masalah dari penelitian ini, yaitu:

1. Bagaimana menerapkan sistem tata letak material untuk membantu persediaan material menggunakan metode asosiasi algoritma *apriori*?
2. Bagaimana mengetahui hasil dari perhitungan manual dengan perbandingan menggunakan *software RapidMiner*?
3. Bagaimana melakukan inventori material yang benar?
4. Bagaimana hasil penerapan algoritma *apriori* untuk menentukan tata letak material pada PT BATAM CYCLECT?

1.5 Tujuan Penelitian

Tujuan dalam penelitian merupakan jawaban atau sasaran yang ingin dicapai dalam sebuah penelitian. Adapun tujuan penelitian ini yaitu:

1. Mengolah data yang ada dan menyusun tata letak material menurut kelompoknya menggunakan algoritma *apriori*.
2. Untuk mengetahui hasil dari perhitungan manual dengan perbandingan menggunakan *software RapidMiner*.
3. Membantu storeman mempersingkat pencarian material dan ruangan didalam gudang lebih teratur dan Membantu storeman melakukan pengecekan material.

4. Untuk menerapkan hasil dari algoritma *apriori* dalam menentukan tata letak material pada PT BATAM CYCLECT.

1.6 Manfaat Penelitian

Dengan penelitian ini dapat diperoleh manfaat bagi para pembaca. Manfaat pada penelitian ini terbagi menjadi manfaat teoritis dan manfaat praktis, yaitu :

1.6.1 Manfaat Teoritis

1. Penelitian ini dapat memberikan gambaran tentang algoritma *apriori* dalam membantu pengaturan tata letak.
2. Sebagai media sistem yang membantu storeman dan referensi perusahaan yang memiliki gudang.
3. Penelitian ini diharapkan dapat digunakan sebagai referensi tambahan untuk membantu para peneliti dalam penelitian selanjutnya mengenai *Data Mining*.

1.6.2 Manfaat Praktis

1. Bagi Peneliti

Untuk meningkatkan pengetahuan penulis mengenai *Data Mining* terutama metode asosiasi algoritma *apriori*.

2. Bagi Universitas Putera Batam

Manfaat bagi mahasiswa khususnya di Universitas Putera Batam dapat mengetahui secara luas pemanfaatan *Data Mining* dalam bidang logistik.

3. Bagi PT BATAM CYCLECT.

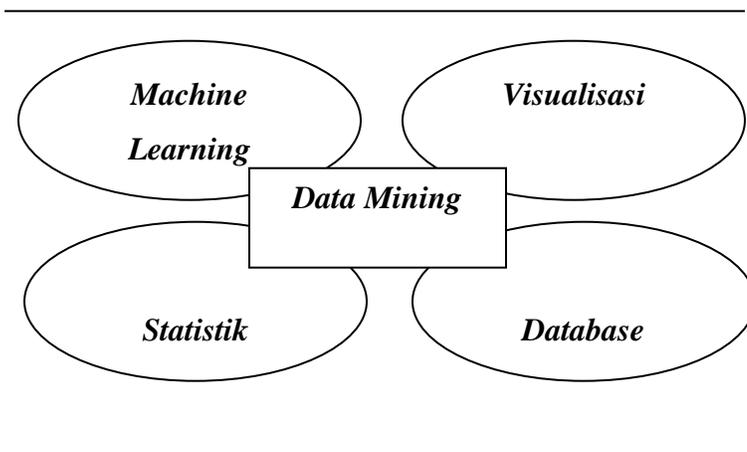
Dapat memberikan informasi mengenai Data Mining dan Membantu untuk mengambil keputusan dalam pengaturan tata letak material.

BAB II

KAJIAN PUSTAKA

2.1 *Knowledge Discovery In Database (KDD)*

Kecerdasan buatan lebih bersifat permanen dari pada kecerdasan alami karena kecerdasan alami akan cepat mengalami perubahan. Kecerdasan buatan konsisten dan murah. Istilah lain *data mining* yaitu KDD atau *knowledge discovery in database*. Tujuan KDD yaitu untuk memanfaatkan data untuk diolah menjadi suatu informasi yang berguna yang diilustrasikan (Tana et al., 2018) sebagai berikut :



Gambar 2. 1 Akar Bidang Ilmu
Sumber: (Peneliti, 2020)

Data mining memiliki empat akar bidang ilmu, yaitu :

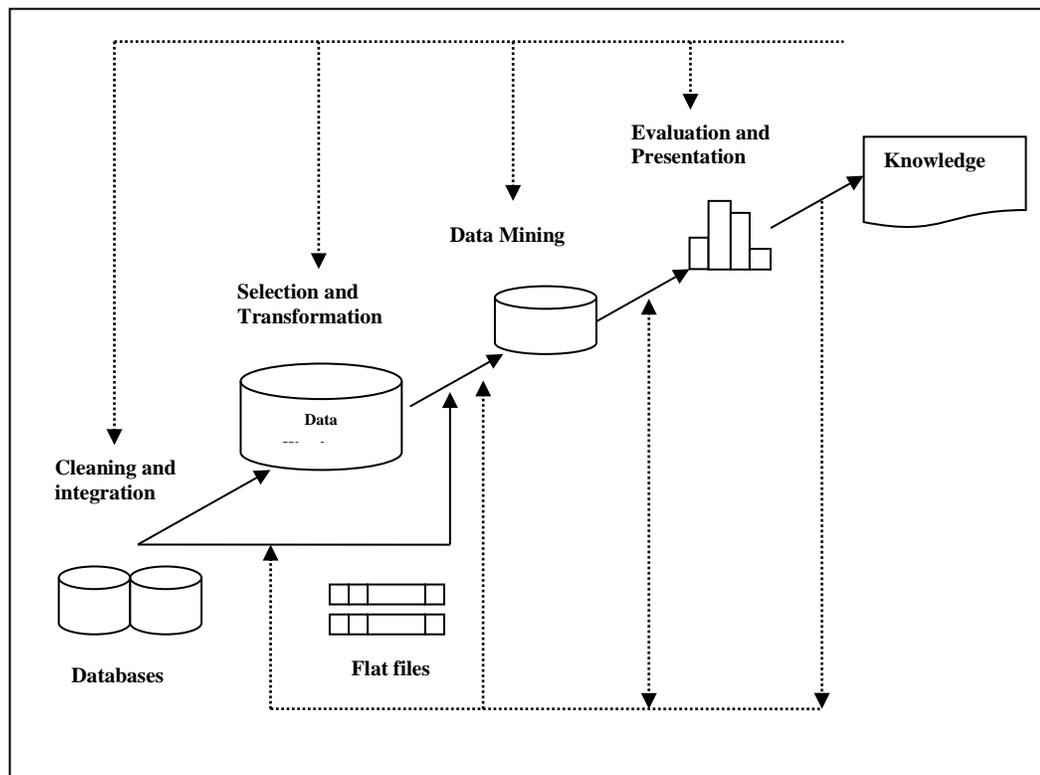
1. Statistik, yaitu bidang ilmu akar tertua, tidak akan ada *data mining* tanpa adanya bidang ini.

2. Kecerdasan buatan, yaitu teori yang dibangun berdasarkan penalaran manusia. Salah satu cabang ilmunya adalah *machine learning*. *Machine learning* adalah disiplin ilmu yang sangat penting dalam membangun data mining, dimana sebuah sistem komputer belajar dengan proses latihan.
3. Visualisasi, yaitu ciri khas dalam data mining karena pencarian pola data mining yaitu pola asosiasi dan pola sekuensial. Pada dasarnya pengenalan pola dalam data mining terbatas pada pola dari *database*.
4. *Database* (Sistem Basis Data), yaitu menyediakan informasi berupa data yang dapat digali dengan menggunakan metode-metode tertentu.

KDD adalah kegiatan yang melibatkan pengumpulan, pemakaian data, historis, untuk menemukan keteraturan, pola atau hubungan dalam set data yang berukuran besar. KDD adalah suatu proses pengganti data mentah menjadi lebih bermanfaat dan sebagai sumber pengetahuan.

2.1.1 Proses *Knowledge Discovery In Database*

Data mining dibagi menjadi beberapa tahapan dimulai dari data mentah sampai dengan pengetahuan atau informasi yang sudah diolah.



Gambar 2. 2 Proses *Knowledge Discovery in Database* (KDD)

Sumber: (Peneliti, 2020)

Tahapan *knowledge discovery in database* (KDD) ada 7, yaitu :

1. Pembersihan data (*Data cleaning*) merupakan proses pembuangan *noise* dan data yang tidak relevan atau konsisten.
2. Integrasi data (*data intergration*) merupakan penggabungan data dari beberapa *database* menjadi satu. Intergrasi dapat dilakukan pada objek yang unit seperti atribut nama, jenis produk, nomor pembeli dan lainnya.
3. Transformasi data (*data transformation*) merupakan data yang sudah dipilih kemudian diubah atau digabungkan kedalam format data mining untuk diproses.

4. Seleksi data (*data selection*) merupakan proses seleksi atau pemilihan data yang relevan terhadap penelitian untuk diterima dari koleksi data yang sudah ada. Sebagai contoh sebuah kasus yang meneliti faktor kecenderungan orang mencari barang dalam kasus analisis tata letak barang.
5. Proses *mining* merupakan proses paling penting karena dilakukan berbagai teknik yang diaplikasikan untuk mendapatkan berbagai pola potensial untuk mendapatkan data yang terselubung.
6. Evaluasi pola (*pattern evaluation*) merupakan tujuan menemukan pola-pola menarik yang sebelumnya telah diidentifikasi berdasarkan *measure* yang telah diberikan. Dalam tahapan ini hasil data mining berupa pola-pola yang unik maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada sudah tercapai, bila belum tercapai maka harus menggunakan teknik data mining yang lain sampai menemukan hipotesa yang tepat.
7. Presentasi pengetahuan (*knowledge presentation*) merupakan penyajian pengetahuan mengenai metode yang digunakan untuk mencapai pengetahuan yang dapat dipergunakan oleh *user*. Penyajian hasil *data mining* dalam bentuk wawasan yang mudah dipahami semua orang.

2.2 Data Mining

2.2.1 Sejarah Data Mining

Tahun 1990-an telah lahir sebuah data di bidang ilmu pengetahuan, bisnis dan pemerintah. Sebelum tahun 1990-an *data mining* dikenal sebagai sub proses dalam cakupan KDD, *data mining* terkait dengan penemuan informasi yang tersembunyi. Sub

data lain yang merupakan bagian dari KDD diantaranya adalah *warehousing*, *data cleaning*, *pre-proses*, dan lain-lain. Untuk beberapa tujuan KDD dan data mining sering dianggap sinonim, tetapi secara teknis ternyata yang satu merupakan sub proses dari yang lain. Data yang digunakan dalam proses data mining pada mulanya hanya untuk data dalam bentuk tabel mengingat keterbatasan kemampuan komputasi saat itu. Saat ini, dalam konteks data bentuk tabel telah diperlakukan secara khusus dengan teknik tersendiri. Hal ini dapat dilihat dari banyaknya sistem seperti RapidMiner, MATLAB, SPSS dan WEKA yang melakukan penambahan data dari dalam tabel.

Popularitas data mining semakin meningkat secara berkala pada tahun 1990-an, terutama dalam pelaksanaan sejumlah konferensi yang dikhususkan pada kajian tersebut, seperti ACM SIGKDD *annual conference* tahun 1995, European PKDD and Pacific/Asia (PAKDD) pada tahun 1997. IEEE ICDM mulai diadakan pada tahun 2001 sebagai konferensi SIAM yang pertama. Hal ini meningkatkan popularitas data mining yang secara bersamaan didukung oleh kemajuan teknologi, kemampuan CPU dan media yang menyediakan data dalam jumlah besar dan memprosesnya dalam waktu yang lebih cepat. Data yang terseimpan dalam jumlah besar dapat ditambang untuk menentukan pola/perilaku pembelian dalam sebuah bisnis. Popularitas data mining terus berkembang samapai saat ini terlebih dalam hal menambang data dari data yang tidak terstandar.

2.2.2 Pengertian Data Mining

Data Mining merupakan salah satu cabang ilmu komputer yang relatif baru. Dan sampai saat ini masih banyak orang memperdebatkan dalam penempatan data mining di bidang ilmu mana, karena data mining menyangkut database, kecerdasan buatan, statistik dan masih banyak lagi.

Menurut (Anjar Wanto, 2020) Data Mining adalah rangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu database dengan melakukan penggalian pola-pola dari data dengan tujuan untuk memanipulasi data menjadi informasi yang lebih berharga yang diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat dalam *database*.

Proses *Knowledge Discovery in Database* melibatkan hasil proses data mining (proses ekstrak suatu pola data), kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami.

Hal penting yang terkait dengan data mining yaitu:

1. Data Mining merupakan proses otomatis terhadap data yang sudah ada.
2. Data yang diproses berupa data yang berukuran besar.
3. Tujuan data mining adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat atau berguna.

Alasan utama mengapa data mining sangat menarik perhatian industri informasi dalam beberapa tahun belakangan ini adalah karena tersedianya data dalam jumlah yang besar dan semakin besarnya kebutuhan untuk mengubah data tersebut menjadi

informasi dan pengetahuan yang berguna. Kehadiran data mining dilatar belakangi dengan *problem data explosion* yang dialami akhir-akhir ini dimana banyak organisasi telah mengumpulkan data sekian tahun lamanya.

Data Mining biasanya diterapkan dalam 3 langkah (Anjar Wanto, 2020) :

1. Eksplorasi

Melakukan persiapan data dalam jumlah besar, kemudian data disaring atau direduksi sesuai yang dibutuhkan, dan menghapus data – data yang diduplikat sehingga data yang tersisa adalah data yang dapat dipergunakan.

2. Pemodelan atau Identifikasi Pola

Membentuk model statistik dengan tujuan mengevaluasi sehingga memperoleh hasil yang terbaik dan paling akurat. Proses ini membutuhkan waktu yang panjang karena model yang digunakan berbeda pada kumpulan data yang sama dan dilakukan proses berulang - ulang untuk dibandingkan hasilnya.

3. Penerapan

Tahapan akhir ini dilakukan pengujian model yang digunakan terhadap data uji dan data lapangan untuk memperoleh hasil prediksi atau perkiraan yang sesuai dengan yang diharapkan.

223 Manfaat Data Mining

Menurut (Vulandari, 2017), Data Mining merupakan pemanfaatan dari 2 sudut pandang yaitu sudut pandang komersial dan sudut pandang keilmuan.

1. Sudut pandang komersial digunakan untuk menangani ledakan *volume* data, dengan memanfaatkan teknik komputasi dapat digunakan untuk memperoleh hasil informasi yang dibutuhkan berupa asset yang dapat meningkatkan daya saing suatu institusi.

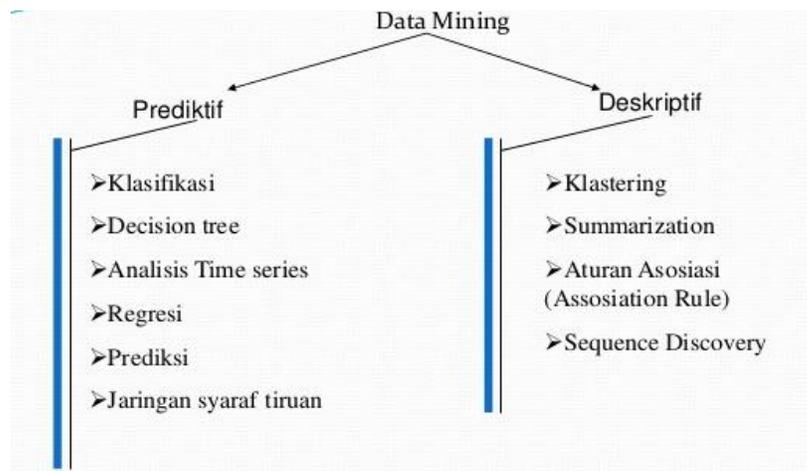
Contoh : Bagaimana mengetahui item barang atau konsumen yang memiliki kesamaan karakteristik ?

2. Sudut pandang keilmuan digunakan untuk menganalisis, memprediksi serta menyimpan data yang bersifat real time dan sangat besar.

Contoh : *Telescope* yang digunsksn untuk memindai langit.

2.2.4 Strategi Data Mining

Menurut (Vulandari, 2017), Data Mining dibagi menjadi 2 kategori yaitu :



Gambar 2. 3 Strategi Data Mining

Sumber: (Peneliti, 2020)

1. Prediktif

Prediktif bertugas untuk memprediksi nilai dari atribut tertentu berdasarkan pada nilai atribut – atribut lain. Atribut yang diprediksi ini menjadi target atau variable tak bebas, sedangkan atribut – atribut yang digunakan untuk mendapatkan prediksi dikenal sebagai *explanatory* atau variable bebas.

Contoh : Perusahaan – perusahaan sangat membutuhkan prediksi calon konsumen yang paling potensial yang akan membeli produknya sehingga dapat menghemat dana iklan perusahaan. Jadi perusahaan membutuhkan teknik – teknik peramalan.

2. Deskriptif

Deskriptif bertugas untuk menurunkan pola – pola (kolesi, trend, cluster, teritori dan anomaly) yang diringkas dalam hubungan sebuah data. Tugas data mining deskriptif merupakan penyelidikan dan seringkali memerlukan teknik *post*.

Contoh : Gudang penyimpanan barang bisa memanfaatkan pola ini untuk meletakkan barang yang memiliki hubungan secara berdekatan untuk meningkatkan proses pencarian barang.

2.3 Metode *Data Mining*

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan (sihombing, 2020), yaitu :

1. *Classification* (Klasifikasi)

Klasifikasi adalah teknik manipulasi data yang telah dikategorikan dan menggunakan hasil untuk memberikan sejumlah aturan. Teknik yang paling populer menggunakan klasifikasi yaitu *Decision Tree* karena mudah untuk diinterpretasi merupakan model prediksi menggunakan struktur pohon atau struktur berhirarki.

2. *Association* (Asosiasi)

Asosiasi adalah teknik untuk mengenali kelakuan dari kejadian khusus atau sebuah proses dimana hubungan asosiasi muncul setiap kejadian. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja'. Contoh asosiasi dalam bisnis dan penelitian adalah menentukan barang yang sering di pesan secara bersamaan dan yang tidak dipesan secara bersamaan.

3. *Clustering* (Pengklusteran)

Pengklusteran adalah teknik analisis pengelompokan berbeda terhadap data. Biasanya teknik ini menggunakan metode *neural network* atau statistik. Contoh pengklusteran adalah mendapatkan kelompok pembeli untuk target pemasaran dari satu produk bagi perusahaan yang tidak memiliki dana pemesanan yang besar.

4. Prediksi

Prediksi adalah teknik yang hampir mirip dengan klasifikasi dan estimasi, tetapi hasil dari nilai prediksi untuk masa yang mendatang. Contoh prediksi adalah prediksi presentasi kenaikan pasien Covid tiga bulan mendatang.

5. Estimasi

Estimasi adalah teknik ini hampir sama dengan teknik klasifikasi, kecuali variable target estimasi lebih kearah numerik dari pada ke kategori. Contoh kasus dari estimasi adalah akan diadakan estimasi sistem kekebalan tubuh pada pasien rumah sakit berdasarkan umur pasien dan gender pasien. Model estimasi yang dihasilkan dapat digunakan untuk penelitian baru kedepan.

6. Deskripsi

Deskripsi adalah teknik untuk menjelaskan pola yang sedang tren terjadi. Hasil data untuk menggambarkan pola dan kecenderungan yang terdapat dalam data.

2.4 Algoritma Data Mining

2.4.1 Association Rules

Association Rules adalah salah satu metode yang bertujuan mencari pola yang sering muncul di antara transaksi, dimana setiap melakukan transaksi terdapat dari beberapa *item* (Fauzy et al., 2016). Analisis *association Rules* adalah salah satu metode data mining yang terkenal pengaplikasinya untuk menganalisis isi keranjang belanja, analisis ini sering disebut sebagai *market basket analysis*.

Definisi – definisi yang terdapat pada *association Rule*, yaitu :

1. I adalah himpunan yang tengah dibicarakan.

Contoh {*Marking Tape, Isolation Tape, ..., Handgloves*}

2. D adalah himpunan seluruh transaksi yang tengah dibicarakan.

Contoh {Transaksi 1, Transaksi 2, ..., Transaksi 15}

3. Proper Subset adalah himpunan bagian murni.

Contoh adalah Himpunan bagian $X = \{x,y,z\}$; Himpunan bagian dari X adalah Himpunan Kosong $\{\}$, Himpunan 1 unsur $\{x\}, \{y\}, \{z\}$ dan himpunan 2 atau lebih unsur.

4. *Itemset* adalah himpunan 1 unsur dan himpunan 2 unsur.

Contoh adalah ada suatu himpunan $X = \{x,y,z\}$ terdiri dari $\{x\}; \{y\}; \{z\}; \{x,y\}; \{x,z\}; \{y,z\}$

5. *K-itemset* adalah *itemset* yang terdiri dari k buah *item* yang ada pada i, intinya k adalah jumlah unsur yang terdapat pada suatu himpunan. Contoh adalah 3-*itemset* adalah yang bersifat 2 unsur.

6. *Itemset* frekuensi adalah jumlah transaksi di i yang mengandung jumlah *itemset* tertentu. Intinya adalah jumlah transaksi yang memesan suatu *itemset*.

7. *Frequent itemset* adalah *itemset* yang muncul sekurang-kurangnya sekali di D. Kata sekali biasanya disimbolkan dengan lambang \emptyset . \emptyset merupakan batas *minimum* dalam suatu transaksi.

F_k adalah himpunan semua *frequent itemset* yang terdiri dari k-item.

Aturan asosiasi dikatakan penting atau tidaknya dapat diketahui dengan dua parameter, *support* (nilai penunjang) yaitu presentase kombinasi *database* sedangkan *confidence* (nilai kepastian) yaitu untuk memperkuat hubungan antara item dalam aturan asosiasi. Analisis asosiasi merupakan suatu proses untuk menemukan semua aturan asosiatif yang telah memenuhi syarat minimum untuk *support* (minimum *support*) dan syarat minimum untuk *confidence* (minimum *confidence*) menurut

(Hasibuan et al., 2017). Khusus salah satu tahapan dari aturan asosiasi yaitu untuk menganalisis pola *frequency* tinggi (*frequent pattern mining*) menarik banyak peneliti untuk menghasilkan nilai yang lebih efisien.

Association Rule yang digunakan adalah “*if ... then...*” atau “jika ... maka...” adalah hasil dari fungsi aturan asosiasi. Dimana istilah yang disebut dengan *antecedent* untuk mewakili bagian dari “*if*” dan *consequent* untuk mewakili bagian “*then*”. Untuk analisis *antecedent* dan *consequent* adalah kelompok yang tidak memiliki hubungan secara bersama.

Langkah – langkah algoritma pada *association Rule*, yaitu :

1. Tentukan \emptyset
2. Tentukan semua *frequency itemset*.

Untuk setiap frekuensi *itemset* lakukan hal,yaitu ambil sebuah unsur dan beri nama A, untuk sisanya beri nama AA-A setelah itu masukkan unsur-unsur yang telah diumpamakan kedalam *Rule if(AA-A) then A* untuk langkah ke 3 lakukan untuk semua unsur.

Assocation Rule merupakan sebuah ekspresi implikasi yang terbentuk dari $A \rightarrow B$, dimana A dan B merupakan disjoint *itemset* ($A \cap B = \emptyset$). Tahapan yang dilakukan dengan mencari kombinasi *item* berdasarkan *minimum support* (Fauzy et al., 2016). Nilai *support* sebuah item diperoleh dengan rumus berikut:

Rumus 2. 1 Support 1-Itemset

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

Sumber: (Peneliti, 2020)

Sementara itu nilai support dari 2 item diperoleh dari rumus 2 sebagai berikut:

Rumus 2. 2 Support 2-Itemset

$$\text{Support (A} \cap \text{B)} = \frac{\text{Jumlah transaksi mengandung A dan B}}{\text{Total transaksi}}$$

Sumber: (Peneliti, 2020)

Setelah semua pola *frequency* tinggi, dicarilah aturan asosiatif yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif A → B, dari aturan asosiatif A → B maka diperoleh rumus sebagai berikut :

Rumus 2. 3 Confidence

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

Sumber: (Peneliti, 2020)

2.4.2 Algoritma Apriori

Algoritma *Apriori* pertama kali diperkenalkan oleh Agrawal dan Shrikant tahun 1994 yang berguna untuk menemukan *frequent itemset* pada sekumpulan data. Selain itu dari *Apriori* terdapat beberapa *algoritma association Rule* lainnya (Fitriati, 2016).

Algoritma *Apriori* termasuk jenis teknik *association Rule* pada data mining. Algoritma *Apriori* bertujuan untuk menemukan *frequent itemsets* pada sekumpulan

data. Algoritma *Apriori* merupakan suatu proses untuk menemukan suatu aturan apriori yang memenuhi syarat *minimum support* dan syarat *minimum confidence*. Algoritma *Apriori* diterapkan untuk membantu dalam membentuk kandidat kombinasi *item*, kemudian dilakukan pengujian apakah kombinasi tersebut memenuhi parameter *support* dan *confidence minimum* yang merupakan nilai ambang yang diberikan oleh pengguna.

Menurut (Rizkiana et al., 2019), Algoritma *Apriori* adalah algoritma pengambilan data dengan aturan asosiatif untuk menentukan hubungan asosiatif suatu kombinasi *item*. Cara algoritma bekerja adalah algoritma akan menghasilkan kandidat baru dari *k-itemset* dari *frequent itemset* pada langkah sebelumnya dan menghitung nilai *support k-itemset* tersebut. *Itemset* yang memiliki nilai *support* dibawah dari *minsup* akan dihapus. Algoritma berhenti ketika tidak ada lagi *frequent itemset* baru yang dihasilkan.

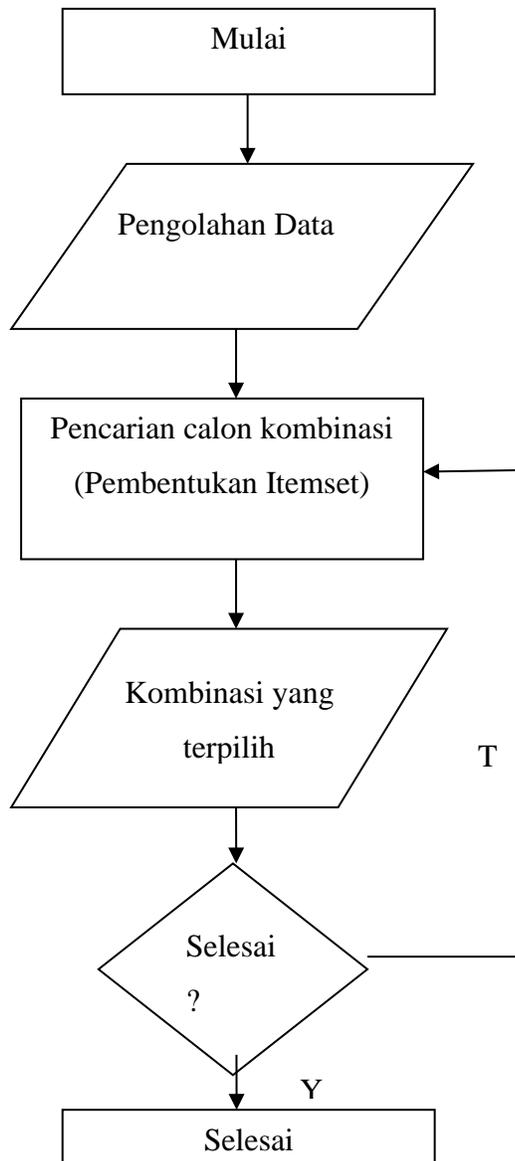
Selanjutnya hasil *frequent itemset* tersebut dihitung *minconf* mengikuti rumus sesuai yang telah ditentukan. *Support* tidak perlu dilihat lagi, karena *generate frequent itemset* didapatkan dari melihat *minsup*. Bila *Rule* yang didapatkan memenuhi batasan yang ditentukan dan batasan itu tinggi, maka *Rule* tersebut tergolong *strong Rules*.

Untuk mengetahui kekuatan aturan asosiatif dapat diukur engan *lift ratio*. *Lift ratio* bertujuan mencari nilai untuk mengukur seberapa besar pentingnya pola yang terbentuk. Maka diperoleh rumus sebagai berikut :

Rumus 2. 4 *Lift Ratio*

$$Lift\ ratio = \frac{Support\ (A \cap B)}{Support\ (A)} \times Support\ (B)$$

Sumber: (Peneliti, 2020)



Gambar 2. 4 Flowchart Algoritma Apriori
Sumber: (Peneliti, 2020)

Langkah awal algoritma *Apriori* yaitu menghitung nilai *support* setiap *item* dalam *database*. Setelah *support* dari *item* ditemukan selanjutnya menentukan

minimum nilai *support*, *item* yang memenuhi syarat nilai minimum dipilih sebagai pola *frequency* tinggi dengan panjang 1 atau disebut dengan *1-itemset*.

Langkah selanjutnya akan menghasilkan *2-itemset* dimana setiap set-nya terdiri dari dua *item*. *Item* yang memenuhi syarat pola *frequency* tinggi pada *1-itemset* dikombinasikan lalu didapatkan calon *frequency 2-itemset* yang selanjutnya dihitung nilai minimum *support* dengan memindai *database*. *Item* yang tidak memenuhi nilai minimum *support* dibuang lalu *item* yang memenuhi nilai minimum *support* dipilih sebagai pola *frequency* tinggi.

Iterasi berhenti jika *item* telah dihitung dan tidak ada kombinasi antara *item* lagi atau tidak ada lagi *support* yang memenuhi nilai minimum *support* yang telah ditetapkan sebelumnya. Kemudian hasil kombinasi dari *k-itemset* yang memenuhi pola *frequency* tinggi dihitung *confidence*.

243 Algoritma C4.5

Algoritma C4.5 dapat membuat pohon keputusan (*Decision Tree*) yang masing-masing internal *node*-nya sebagai atribut pengujian, setiap cabang mewakili output dari pengujian, dan setiap *node* daun (*terminal node*) menentukan *class*. Algoritma C4.5 menggunakan informasi *gain* atau *entropy reduction* untuk memilih pembagian yang optimal. Algoritma C4.5 dapat mengatasi *missing value*, *continue data* dan pruning dari pengembangan algoritma C4.5 (Widayu et al., 2017).

244 Algoritma FP-Growth

FP-Growth adalah algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam kumpulan data. *FP-Growth*

menggunakan pendekatan yang berbeda dari yang digunakan algoritma *apriori* (Triyanto, 2014). *FP-Growth* dikembangkan dari algoritma *apriori*. *FP-Growth* juga dikenal dengan *FP-Tree*, karena menggunakan struktur *FP-Tree* yang dapat langsung diekstrak *frequent itemset* dari susunan *FP-Tree* yang telah dibentuk.

2.5 *Market Basket Analysis*

Market basket analysis merupakan teknik pemodelan dalam data mining berdasar teori dimana anda membeli suatu grup *item*, anda akan memiliki kemungkinan membeli *itemset* yang lain (Fatihatul et al., 2011).

Market basket berdasarkan dari kumpulan *item* yang dibeli oleh konsumen dalam sebuah transaksi. Dalam proses ini, kuantitas dari sebuah *item* tidak mempengaruhi proses analisis. Dalam *market basket analysis* akan dilakukan analisis akumulasi kumpulan transaksi dari sejumlah besar konsumen dalam periode waktu yang telah berlangsung.

Proses ini menganalisis *buying habits* dari para konsumen dengan mengumpulkan hubungan asosiasi antar *item-item* berbeda yang seringkali dibeli oleh konsumen. Hasil proses ini nantinya akan sangat bermanfaat bagi perusahaan retail khususnya toko swalayan, supermarket dan gudang untuk mengembangkan strategi pemasaran dan proses pengambilan keputusan dengan melihat *item-item* berbeda yang seringkali dibeli secara bersamaan.

Beberapa kombinasi *item* yang sering dibeli konsumen yang mudah ditebak, contohnya seringkali konsumen membeli roti tawar dan selai secara bersamaan. Pola

ini sangat biasa terjadi dan mudah ditebak karena roti tawar dan selai memiliki hubungan yang bersamaan (Elisa, 2018). Namun terkadang pola seperti gula dan sabun jarang terpikir karena gula dan sabun tidak memiliki hubungan sama sekali. Namun, melalui *market basket analysis* pola – pola yang tidak terpikir dapat ditemukan dengan mudah sehingga memudahkan pengambilan keputusan dan proses pengembangan strategi bagi perusahaan retail. Masalah seperti kehabisan stok dapat diminimalisir dengan diketahui pola pembelian konsumen melalui *market basket analysis* sehingga diperoleh peningkatan dalam penjualan sebuah perusahaan.

2.6 *Software* Pendukung

Penelitian yang dilakukan menggunakan sebuah *software* utama untuk melakukan pengujian. *Software* tersebut adalah RapidMiner.

2.6.1 *RapidMiner*



Gambar 2. 5 Logo RapidMiner

Rapidminer merupakan salah satu *tool* data mining yang digunakan peneliti, akademis dan pihak lain untuk mempermudah proses pengolahan. RapidMiner adalah pengolahan data yang bertasir dengan analisis teks, mengekstrak pola-pola dari dataset

yang besar dan mengkombinasikan dengan metode statistika, kecerdasan buatan dan database (Faid et al., 2019).

RapidMiner merupakan *software* yang dikembangkan oleh Raif Klinkenberg, Ingo Mierswa, dan Simon Fischer pada tahun 2001 di *Artificial Intelligence Unit* dari *Universitas Of Dortmund* yang sebelumnya dikenal dengan nama YALE (*Yale Another Learning Environment*). Pada tahun 2007 perangkat lunak YALE berubah nama menjadi RapidMiner yang sebelumnya bernama Rapid-I.

2.7 Tata Letak Barang

Tata letak yang teratur akan memberikan kesan dan memudahkan proses pencarian barang bagi konsumen. Hal tersebut yang menjadi salah satu alasan bagi konsumen memilih toko karena memberikan suasana yang nyaman dan teratur. Menempatkan barang dengan teratur dan rapi dapat dilakukan dengan melihat kategori barang sesuai dengan jenis dari barang tersebut. Menempatkan barang-barang yang berkaitan yang sering dibeli atau habis harus menjadi perhatian dari produsen untuk peletakannya pada bagian rak yang mudah terjangkau atau terlihat.

2.8 Penelitian Terdahulu

Penelitian terdahulu digunakan sebagai referensi untuk peneliti dalam melakukan penelitian untuk menentukan tata letak menggunakan data mining.

Ada beberapa jurnal yang peneliti jadikan sebagai referensi sebagai berikut:

1. Erlin Elisa dari Program Studi Sistem Informasi, Fakultas Teknik, Universitas Putera Batam [2018] Jurnal Resti : Rekayasa Sistem dan Teknologi Informasi, e-ISSN: 2580-0760, Vol. 2, No. 2, dengan judul “**MARKET BASKET ANALYSIS PADA MINI MARKET AYU DENGAN ALGORITMA APRIORI**” (Elisa, 2018). Manfaatkan pola dan strategi dalam proses transaksi itu sangat penting. Tujuan penelitian ini untuk menerapkan algoritma *Apriori* pada mini market dalam menentukan penjualan atau barang konsumsi yang sering dibeli atau habis terjual. Dari hasil penerapan algoritma *Apriori* dapat dicari pola asosiasi berdasarkan pola pembelian yang dilakukan konsumen, sehingga bisa ditemukan *item-item* barang yang sering dibeli secara bersamaan. Setelah menerapkan algoritma *Apriori* dalam penjualan terdapat nilai *support* dan *confidence* tertinggi yaitu minyak dan susu dengan nilai *support* 42,85% dan *confidence* 85,71%. Hasil tersebut diuji menggunakan aplikasi Tanagara.
2. Robi Yanto, Riri Khoiriah dari Program Studi Sistem Informasi, STMIK Bina Nusantara Lubuklingau [2015] Citee Journal, e-ISSN: 2354-5771, Vol. 2, No.2, dengan judul “**IMPLEMENTASI DATA MINING DENGAN METODE ALGORITMA APRIORI DALAM MENENTUKAN POLA PEMBELIAN OBAT**” (Yanto & Khoiriah, 2015). Disimpulkan tujuan dari penelitian ini adalah algoritma *Apriori* dapat menentukan pola ketersediaan obat dan tata letak obat oleh konsumen. Algoritma *Apriori* memudahkan kita mengetahui letak suatu obat dari hubungan 2 itemset obat. Dalam dunia medis untuk mengetahui kebiasaan konsumen maka diterapkan algoritma *Apriori* untuk menentukan

parameter *support* dan *confidence* minimum yang merupakan nilai ambang yang diberikan konsumen. Maka disimpulkan jika membeli amoxicillin maka membeli obat asamefenamat dengan *confidence* 75 % dengan menggunakan perhitungan sistem yang telah dibangun juga didapatkan hasil yang sama. Dengan mengetahui nilai dari sebuah obat dapat membantu pihak karyawan dalam mengatur letak obat yang terdiri dari 2 *itemset* secara berdekatan dan memudahkan karyawan mencari keberadaan obat.

3. Helmanatun Nisa Wulandari, Nur Wijayaning Rahayu dari Prodi Teknik Informatika, Universitas Islam Indonesia [2014] Seminar Nasional Aplikasi Teknologi Informasi (SNATI), e-ISSN: 1907-5022 dengan judul **“PEMANFAATAN ALGORITMA APRIORI UNTUK PERANCANGAN ULANG TATA LETAK BARANG DI TOKO BUSANA”** (Wulandari & Rahayu, 2014). Penelitian ini bertujuan untuk menerapkan algoritma Apriori untuk menganalisis pola belanja konsumen di toko busana muslimah. Dari pola konsumen tersebut maka dihasilkan itemset yang berkaitan atau dibeli bersamaan, maka disimpulkan barang tersebut harus diletakkan dalam area yang berdekatan. Hasil pengujian tersebut dilakukan pengujian dengan aplikasi WEKA.
4. Nugroho Wandu, Rully A. Hendrawan, Ahmad Mukhlason dari Program Studi Sistem Informasi, Institut Teknologi Sepuluh Nopember [2012] Jurnal Teknik ITS, Vol.1, e-ISSN: 2301-9271 dengan judul **“ PENGEMBANGAN SISTEM REKOMENDASI PENELUSURAN BUKU DENGAN PENGGALIAN**

ASSOCIATION RULE MENGGUNAKAN ALGORITMA APRIORI (STUDI KASUS BADAN PERPUSTAKAN DAN KEARSIPAN PROVINSI JAWA TIMUR)” (Wulandari & Rahayu, 2014). Memanfaatkan data kunjungan perpustakaan untuk penempatan buku dapat mengali informasi tentang buku-buku apa yang sering dipinjam oleh pengunjung dan keterkaitan antar masing-masing peminjaman hingga melakukan penyusunan dan tata letak buku. Solusi dalam permasalahan ini adalah data mining dengan metode *association Rules* dan algoritma *apriori* dari metode dan algoritma tersebut menghasilkan transaksi peminjaman buku dengan *strong association* antar buku dalam mendapatkan rekomendasi buku lain ketika pengguna melihat rincian buku yang dipilih atau yang akan dipinjam. Dari hasil penelitian ini ditemukan bahwa semakin besar *minimum support* dan *minimum confidence*, semakin sedikit waktu yang dibutuhkan untuk menghasilkan rekomendasi serta semakin sedikit rekomendasi yang diberikan, namun rekomendasi yang diberikan berasal dari transaksi yang sering muncul.

5. Apurwa Sahu, Mradul Dhakar, Pushpi Rani dari Department of CSE, ITM University Gwalior, India [2015] *International Journal of Computer Science & Communication (IJCSC)*, Vol. 6 Issue 2, ISSN: 0973-7391, dengan judul **“COMPARATIVE ANALYSIS OF APRIORI ALGORITMA BASED ON ASSOCIATION RULE”** (Patil & Applications, 2018). Penelitian ini bertujuan peningkatan algoritma *apriori* dengan perancangan pendekatan alternatif untuk penambangan aturan asosiasi dan mengurangi kompleksitas waktu. Maka untuk

mengatasi peningkatan data dari waktu ke waktu atau menangani ledakan kuantitas data dibutuhkan pendekatan dari data mining tersebut dengan bantuan algoritma *apriori*.

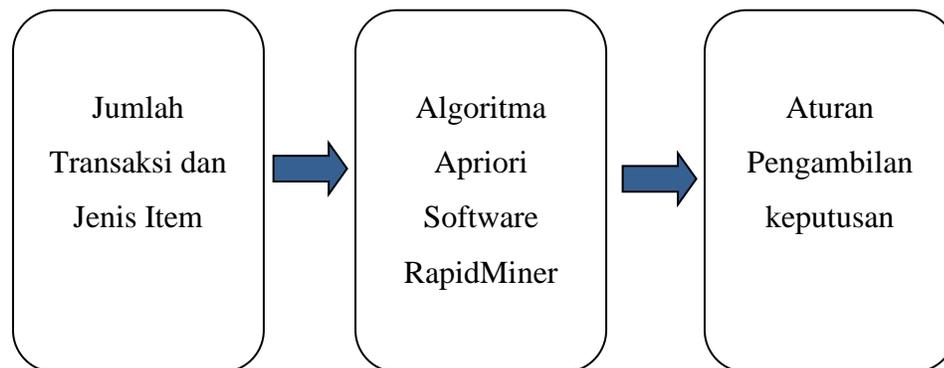
6. Pratibha Mandave, Megha Mane, Prof. Sharada Patil, dari *Department Of Master Of Computer Application*, India [2013] *International Journal Of Latest Trends In Engineering And Tecnology (IJLTET)*, Vol.3 Issue 2, ISSN: 2278-621X, dengan judul “**DATA MINING USING ASSOCIATION RULE BASED ON APRIORI ALGORITHM AND IMPROVED APPROACH WITH ILLUSTRATION**” (Sahu et al., 2015). Dari penelitian ini disimpulkan algoritma *apriori* berguna dan efisien dalam bidang bisnis dan industri di kehidupan sehari-hari. Algoritma *apriori* dapat menghasilkan banyak *itemset*. Frekuensi tinggi dalam sebuah transaksi menghasilkan set *k-itemset*. Jika k lebih besar dari transaksi T yaitu dihasilkan oleh $(k-1)$ *itemset* sesuai sifat algoritma *apriori*.
7. Friskila Parhusip dari Program Studi Sistem Informasi, STIKOM Tunas Bangsa Pematangsiantar [2019] *METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi*, Vol.3, No.1, p-ISSN: 2598-8565, e-ISSN: 2620-4339, dengan judul “**PEMANFAATAN DATA MINING PENEMPATAN BUKU PERPUSTAKAN MENGGUNAKAN METODE ASSOCIATION RULE**” (Parhusip, 2019). Penelitian ini bertujuan untuk menerapkan untuk menerapkan *Data Mining* sebagai pengaturan tata letak di perpustakaan dengan metode *association Rule*. Dengan memanfaatkan dari transaksi peminjam buku diperoleh

kombinasi untuk menyusun tata letak dengan menghubungkan hubungan antar variable yang ada didalam suatu data transaksi.

2.9 Kerangka Pemikiran

Kerangka pemikiran adalah alur pikiran yang logis dalam menghasilkan suatu kesimpulan. Dengan menggunakan Data Mining Algoritma Apriori dapat menentukan Tata Letak Atau *layout* di PT BATAM CYCLECT yang dibutuhkan tahapan yang harus dilakukan untuk mendapatkan hasil rancangan yang dapat berfungsi dengan baik.

Flowchart tersebut sebagai berikut:



Gambar 2. 6 Tahapan Kerangka Pemikiran

Sumber: (Peneliti, 2020)

Dari flowchart diatas dapat diketahui bahwa langkah awal dari penelitian ini yaitu:

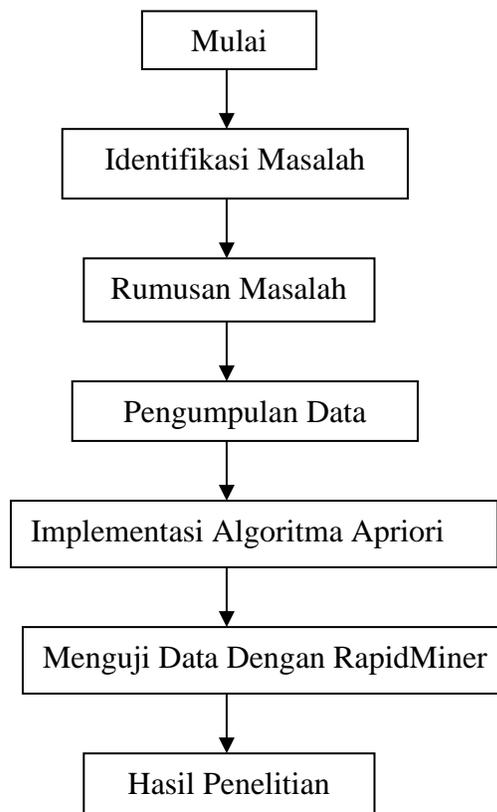
1. Dikumpulkan dan dilakukan *cleansing data*, *transformation* untuk menentukan pola konsumen dari kebiasaan konsumen dalam meminta barang yang dibutuhkan. Tahapan tersebut didapatkan variable masukan berupa jumlah transaksi dan jenis item.
2. Tahapan ini adalah proses pengolahan data menggunakan metode aturan asosiasi algoritma apriori. Kemudian dilakukan pengujian terhadap data-data yang sudah diperoleh melalui *software* RapidMiner.
3. Hasil yang diolah dari penelitian ini berupa aturan untuk pengambilan keputusan berupa aturan asosiasi tertinggi yang dijadikan acuan dalam merekomendasikan tata letak barang.

BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Gambaran penting yang harus dimulai dari awal proses penelitian hingga mendapatkan hasil akhir, sehingga dapat di tarik kesimpulan dari seluruh proses penelitian adalah desain penelitian.



Gambar 3. 1 Desain Penelitian
Sumber: (Peneliti,2020)

Adapun uraian desain penelitian yang dilakukan dalam penelitian ini sesuai *flowchart* desain penelitian diatas, yaitu:

1. Identifikasi Masalah

Identifikasi masalah adalah cara menentukan permasalahan yang akan dibahas dalam sebuah penelitian. Memahami ruang lingkup masalah dan mengetahui latar belakang yang dibahas serta mendefinisikan batasan masalah untuk mendapatkan suatu solusi dari sebuah masalah. Dalam penelitian ini membahas tentang Tata Letak Barang, bagaimana *storeman* menyusun *layout* barang, dengan bantuan aturan asosiasi dan algoritma *apriori* yang memudahkan pencarian barang karena *layout* yang disusun sesuai dengan kategori dan jenis barang.

2. Rumusan Masalah

Setelah mengidentifikasi masalah, kemudian tahap selanjutnya yaitu menjabarkan masalah-masalah yang lebih *detail* lagi.

3. Pengumpulan Data

Agar penelitian ini dapat terstruktur dengan jelas oleh sebab itu peneliti melakukan pengumpulan data dengan 2 cara, yaitu:

- a. Melakukan pengamatan langsung ke PT BATAM CYCLECT yang akan menjadi objek penelitian sehingga masalah yang terjadi dapat peneliti ketahui secara jelas dan melakukan dokumentasi untuk mendapatkan laporan kegiatan penelitian, seperti mengambil gambar dan sebagainya.
- b. Studi literatur yang dilakukan peneliti adalah mempelajari tentang Data Mining, Metode *Association Rule*, dan Algoritma *Apriori* dari jurnal dan beberapa buku

serta artikel yang ada di internet yang berkaitan dengan masalah penelitian tersebut.

4. Implementasi Algoritma *Apriori*

Implementasi Algoritma *Apriori* merupakan tahap selanjutnya dimana setelah masalah yang akan dianalisa ditemukan, selanjutnya akan diolah dengan metode algoritma *apriori* sehingga akan diperoleh nilai *support* dan *confidence* dan menghasilkan pembentukan *strong association* untuk merekomendasikan penyusunan tata letak material.

5. Menguji Data dengan RapidMiner

Setelah tahap implementasi algoritma *apriori*, data yang telah diolah menggunakan algoritma *apriori* akan diuji menggunakan *software* data mining yaitu, RapidMiner untuk memperoleh hasil analisa yang lebih akurat dan melakukan perbandingan dengan perhitungan manual.

6. Hasil Penelitian

Tahapan penutup dari sebuah penelitian yaitu menyimpulkan hasil dari pengolahan data dan pengujian yang telah di proses dan membuat visualisasi untuk penyusunan tata letak material dari hasil yang diolah.

3.2 Teknik Pengumpulan Data

Pengumpulan data adalah langkah yang dilakukan peneliti untuk mendapatkan data dari masalah dalam penelitian. Data mempunyai fungsi bagi peneliti untuk proses

pengolahan menjadi informasi yang berguna. Terdapat beberapa teknik pengumpulan data yang digunakan, yaitu:

1. Observasi merupakan metode yang harus dilakukan ketika melakukan suatu penelitian. Dengan mengumpulkan data secara langsung ke lokasi penelitian. Pengamatan dilakukan untuk mengumpulkan data transaksi barang pada sistem gudang dan melakukan pengamatan terhadap penempatan barang-barang di gudang PT BATAM CYCLECT.
2. Studi *literature* yaitu mempelajari data mining khususnya metode *association Rule* dan algoritma *apriori*, untuk membantu mempermudah peneliti dalam melakukan pengolahan data bersumber dari buku dan jurnal serta artikel yang berasal dari internet.
3. Dokumentasi dilakukan dalam penelitian sebagai bukti peneliti memperoleh data langsung dari tempat penelitian di PT BATAM CYCLECT. Hasil dokumentasi penelitian berupa laporan data mentah dan foto kegiatan penelitian.

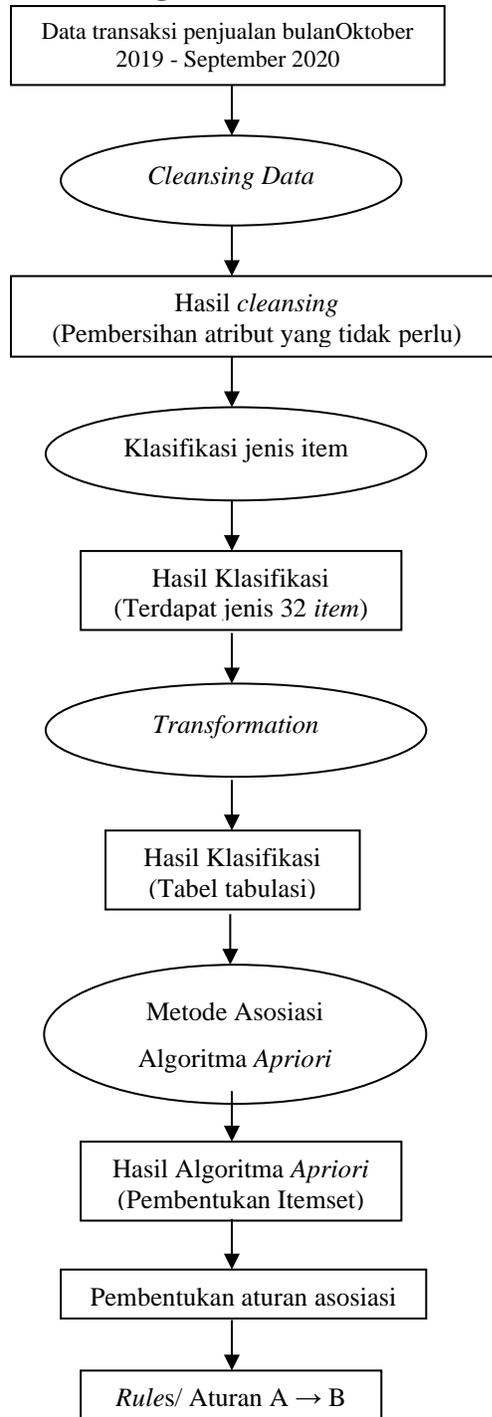
3.3 Operasional Variable

Operasional variable digunakan dalam menentukan suatu atribut atau nilai yang terkait didalam penelitian. PT BATAM CYCLECT adalah perusahaan yang bergerak dibidang *repair* kapal.

Dari pengumpulan data, dilakukan tahap *cleansing* data didapatkan atribut yang akan dijadikan variable dalam penelitian. Atribut yang berperan sebagai variable dalam penelitian ini, yaitu:

1. Nomor transaksi digunakan untuk mengetahui total keseluruhan transaksi yang akan diolah menggunakan metode asosiasi algoritma *apriori*.
2. Jenis *item* digunakan untuk menentukan *item* mana yang memiliki pola frekuensi tinggi dan mempunyai kombinasi antar *item*.

3.4 Metode Analisis dan Rancangan Sistem



Gambar 3. 2 Metode Analisis Rancangan Sistem
Sumber: (Peneliti, 2020)

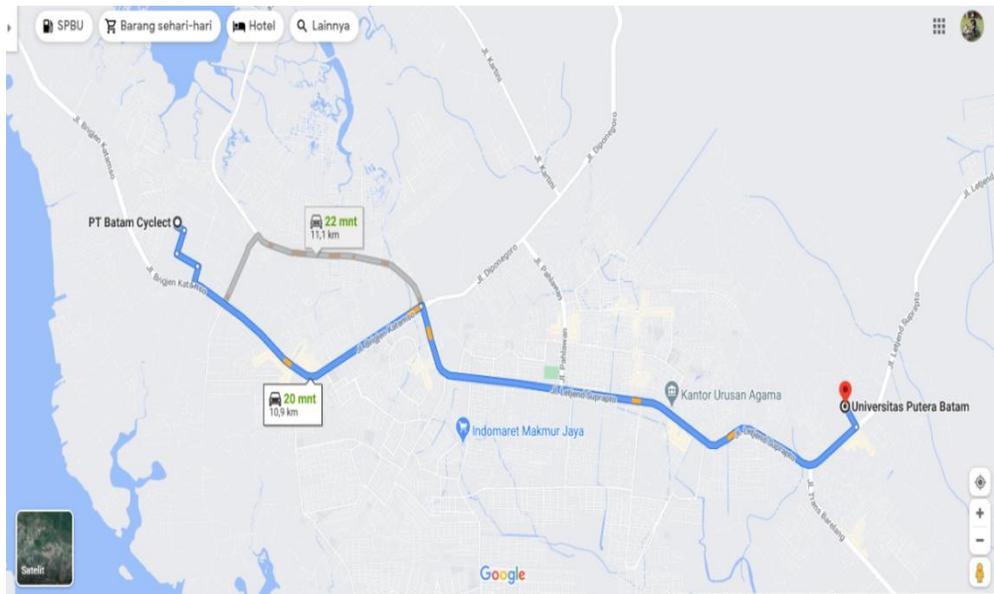
Adapun uraian gambar 3.2 yang dilakukan dalam penelitian ini sesuai *flowchart* metode analisis rancangan sistem diatas, yaitu:

1. Pengumpulan data transaksi order material bulan oktober 2019 sampai dengan september 2020.
2. Tahapan selanjutnya yaitu *cleansing* data (pembersihan data) untuk membersihkan atribut-atribut data yang diperlukan dalam pengolahan data. Setelah data dibersihkan dan data yang tidak diperlukan dibuang dilakukan klasifikasi *item* yang terdiri dari 32 jenis *item*.
3. Selanjutnya dilakukan tahapan klasifikasi data yang berguna untuk mengelompokkan jenis *item* berdasarkan fungsionalitas.
4. Hasil dari klasifikasi data kemudian dilakukan transformasi data yaitu merubah dengan membuat table tabulasi data bilangan biner.
5. Table tabulasi data diproses dengan menggunakan metode *association rule* algoritma *apriori*.
6. Hasil dari proses algoritma *apriori* tersebut terbentuklah *itemset* dan terbentuk aturan asosiasi yang hasil akhir-nya berupa aturan dengan menjabarkan “jika-maka”.

3.5 Lokasi dan Jadwal Penelitian

3.5.1 Lokasi Penelitian

Penelitian ini dilaksanakan di PT BATAM CYCLECT, beralamat di Kawasan Bintang Industri II Blok D1 No. 615-617, Tanjung Uncang.



Gambar 3. 3 Lokasi Penelitian
Sumber: (Peneliti, 2020)

3.5.2 Jadwal Penelitian

Penelitian ini membutuhkan waktu lebih kurang 5 bulan sejak peneliti menemukan permasalahan di PT BATAM CYCLECT dan menemukan solusi yang tepat dari masalah tersebut.

Tabel 3. 1 Jadwal Pelaksanaan Penelitian

NO	Kegiatan	Waktu Pelaksanaan													
		Sep 2020		Okt 2020				Nov 2020		Des 2020			Jan 2021		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Pengajuan judul dan objek penelitian	■													
2	Pengajuan Bab I		■	■	■										
3	Pengajuan Bab II					■	■								
4	Pengajuan Bab III							■	■						
5	Pengolahan Data								■	■	■				
6	Pengajuan Bab IV											■	■		
7	Pengajuan Bab V													■	
8	Pengumpulan Skripsi														■

Sumber:(Peneliti,2020)