

***DATA MINING MEMPREDIKSI JUMLAH  
PENJUALAN TIKET MENGGUNAKAN METODE  
AUTOREGRESSIVE INTEGRATED MOVING  
AVERAGE***

**SKRIPSI**



**Oleh:**

**Cindy Butar-Butar**

**150210260**

**FAKULTAS TEKNIK DAN KOMPUTER  
PROGRAM STUDI TEKNIK INFORMATIKA  
UNIVERSITAS PUTERA BATAM  
TAHUN 2019**

***DATA MINING MEMPREDIKSI JUMLAH  
PENJUALAN TIKET MENGGUNAKAN METODE  
AUTOREGRESSIVE INTEGRATED MOVING  
AVERAGE***

**SKRIPSI**

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



**Oleh:**

**Cindy Butar Butar**

**150210260**

**FAKULTAS TEKNIK DAN KOMPUTER  
PROGRAM STUDI TEKNIK INFORMATIKA  
UNIVERSITAS PUTERA BATAM  
TAHUN 2019**

## **PERNYATAAN**

Dengan ini saya menyatakan bahwa:

1. Skripsi ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (sarjana, dan/atau magister), baik di Universitas Putera Batam maupun di perguruan tinggi lain.
2. Skripsi ini adalah murni gagasan, rumusan, dan penelitian saya sendiri, tanpa bantuan pihak lain, kecuali arahan pembimbing.
3. Dalam skripsi ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
4. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di perguruan tinggi.

Batam, 05 September 2019

Yang membuat pernyataan,

Cindy Butar Butar

150210260

***DATA MINING* MEMPREDIKSI JUMLAH  
PENJUALAN TIKET MENGGUNAKAN METODE  
AUTOREGRESSIVE INTEGRATED MOVING  
AVERAGE**

Oleh  
**Cindy Butar Butar**  
150210260

**SKRIPSI**

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

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

**Batam, 05 September 2019**

**Sunarsan Sitohang,S.Kom., M.TI  
Pembimbing**

## **ABSTRAK**

PT Great Seasons merupakan sebuah perusahaan yang bergerak dalam bidang biro jasa penjualan tiket pesawat, perusahaan ini menjual tiket pesawat dari berbagai maskapai penerbangan. Penjualan tiket pada perusahaan ini mengalami penjualan yang tidak stabil, Oleh karena itu perlu dilakukan sebuah penelitian untuk mengetahui prediksi penjualan dengan teknik Data Mining. Data mining adalah sebuah proses yang dilakukan untuk menentukan pola dari sekelompok data dalam mengambil sebuah keputusan. Adapun metode data mining yang dipakai adalah metode Autoregressive Integrated Moving Average (ARIMA). ARIMA atau yang sering disebut metode runtutan waktu yaitu merupakan teknik dalam mencari sebuah pola yang paling sesuai dalam beberapa kumpulan data dengan identifikasi model, menentukan model terbaik dengan uji statistika dan prediksi untuk waktu yang akan datang. Untuk mendukung proses pengembangan penjualan dengan mengaplikasikan metode ARIMA dalam memprediksi peningkatan penjualan tiket di PT.Great Seasons Travel pada 2019 maka proses data dilakukan dengan melihat data hasil penjualan pada bulan Mei 2018 sampai Oktober 2018. Data yang digunakan akan diproses dengan menggunakan aplikasi Eviews untuk mendapatkan hasil prediksi serta mengetahui langkah apa yang akan dilakukan selanjutnya untuk lebih meningkatkan jumlah penjualan tiket dimasa yang akan datang. Hasil yang diperoleh dari penelitian tersebut yaitu bahwa tingkat penjualan tiket di tahun yang akan datang terprediksi mengalami kestabilan sekitar 95% .

**Kata Kunci :** Penjualan tiket pesawat, Data Mining, ARIMA

## ABSTRACT

PT Great Seasons is a company that sells airline ticket sales, the company sells airline ticket from multiple airlines. The ticket sales of this company are having an unstable sale, therefore a research need to be made to determine predictions of a sale using *Data Mining* techniques. *Data Mining* is a process conducted to determine the pattern a group of data to take a decision and method used is *Autoregressive Integrated Moving Average (ARIMA)* method. Arima commonly called the time building method, is the technique for finding the most suitable pattern in some data collection with model identification, determining the best model for the statistician test and predictability for the future. To support the development of sales by arima's method in ticket sales at PT Great Seasons 2019, the data process is performed by looking at sales data in may 2018 through October 2018. The data used will be processed using the Eviews application to get predictions results and figure out what moves it will take to increase the number of tickets in future sales. The result of the study is that ticket sales in the coming year have been predicted to increase

*Keywords : Airline ticket sales, Data Mining, ARIMA*

## **KATA PENGANTAR**

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

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati.

Dengan segala keterbatasan, penulis menyadari pula bahwa skripsi ini takkan terwujud tanpa bantuan, bimbingan, dan dorongan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

1. Ibu Dr. Nur Elfi Husda S.Kom M.Kom selaku Rektor Universitas Putera Batam.
2. Bapak Andi Maslan S.T M.Si Ketua Program Studi teknik informatika
3. Bapak Sunarsan Sitohang S.Kom.,M.TI selaku pembimbing Skripsi dan pembimbing akademik pada Program Studi Teknik Informatika Universitas Putera Batam.
4. Dosen dan Staff Universitas Putera Batam
5. Orang Tua yang telah mendukung baik secara materi ataupun non materi.
6. Saudara kakak dan adik yang telah mendukung baik secara materi ataupun nonmateri.
7. Amrin Simanjuntak S.kom, Nurmanti, Meliani, selaku sahabat yang yang membantu dan mendukung secara materi ataupun nonmateri.
8. Teman-teman seperjuangan dan rekan kerja

Semoga Tuhan membalas kebaikan dan selalu mencurahkan berkah serta anugerahNYA, Amin.

Batam, 05 September 2019

Cindy Butar Butar

## DAFTAR ISI

<b>HALAMAN SAMPUL DEPAN</b>	
<b>HALAMAN JUDUL</b> .....	<b>ii</b>
<b>HALAMAN PERNYATAAN</b> .....	<b>ii</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>viii</b>
<b>DAFTAR TABEL</b> .....	<b>x</b>
<b>DAFTAR GAMBAR</b> .....	<b>xi</b>
<b>DAFTAR RUMUS</b> .....	<b>xii</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang Masalah.....	1
1.2 Identifikasi Masalah .....	3
1.3 Batasan Masalah.....	4
1.4 Rumusan Masalah .....	4
1.5 Tujuan Penelitian .....	4
1.6 Manfaat Penelitian .....	5
<b>BAB II KAJIAN PUSTAKA</b> .....	<b>7</b>
2.1 <i>Knowledge Discovery In Database (KDD)</i> .....	7
2.2 <i>Data Mining</i> .....	9
2.2.1 Definisi <i>Data Mining</i> .....	9
2.2.2 Pengelompokan <i>Data Mining</i> .....	10
2.3.1 <i>Time Series Analysis</i> .....	12
2.4 Pengertian Penjualan .....	14
2.4.1 Faktor Pengaruh Penjualan .....	15
2.5 Metode <i>Autoregressive Integrated Moving Average (ARIMA)</i> .....	15
2.5.1 Klasifikasi Model ARIMA.....	16
2.5.2 Mengidentifikasi Model .....	20
2.6 <i>Software</i> Pendukung .....	23
2.7 Penelitian Terdahulu .....	23
<b>BAB III METODE PENELITIAN</b> .....	<b>30</b>

3.1 Desain Penelitian (Arsitektur Sistem).....	30
3.2 Teknik Pengumpulan Data.....	32
3.3 Operasional Variabel.....	33
3.4 Metode Analisis Dan Rancangan Sistem ( <i>Data Mining</i> ).....	34
3.5 Lokasi Dan Jadwal Penelitian .....	37
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>39</b>
4.1 Analisis Data .....	39
4.1.1 Pengumpulan Data .....	39
4.1.2 Data Transaksi Penjualan.....	39
4.2 Proses <i>Import</i> Data.....	41
4.2.1 Uji Stasioneritas .....	42
<b>BAB V KESIMPULAN DAN SARAN .....</b>	<b>54</b>
5.1 Kesimpulan .....	54
5.2 Saran.....	54
<b>DAFTAR PUSTAKA .....</b>	<b>54</b>
<b>RIWAYAT HIDUP .....</b>	<b>55</b>
<b>SURAT KETERANGAN PENELITIAN.....</b>	<b>56</b>
<b>LAMPIRAN.....</b>	<b>57</b>

## **DAFTAR TABEL**

Tabel 3.1 Jadwal Penelitian.....	38
Tabel 4.1 Data Transaksi Penjualan.....	40
Tabel 4.2 Tabulasi Data ARIMA Pada Eviews 9 .....	40
Tabel 4.3 Perbandingan Model .....	48
Tabel 4.4 Hasil Forecasting.....	50

## DAFTAR GAMBAR

Gambar 2.1 <i>Knowledge Discovery In Database (KDD)</i> .....	7
Gambar 2.2 Logo Eviews 9.....	23
Gambar 2.3 Kerangka pemikiran .....	29
Gambar 3.1 Desain Penelitian.....	30
Gambar 3.2 Flowchart Of Document Analisis Kebutuhan Sistem .....	35
Gambar 4.1 Halaman depan Eviews 9 .....	42
Gambar 4.2 Hasil Uji stasioneritas.....	43
Gambar 4.3 Hasil Uji autokorelasi dan plot autokorelasi parsial.....	44
Gambar 4.4 Uji model AR (1).....	44
Gambar 4.5 Hasil ARCH tes AR ( 1).....	45
Gambar 4.6 Hasil Uji MA ( 1 ) .....	45
Gambar 4.7 Hasil Uji ARCH MA ( 1 ) .....	46
Gambar 4.8 Hasil Uji ARIMA ( 1 ) .....	47
Gambar 4.9 Hasil Uji ARCH ARIMA ( 1 ) .....	47
Gambar 4.10 Hasil Lag 1 sampai 12.....	48
Gambar 4.11 Graph Hasil Uji Normalitas.....	49
Gambar 4.12 Graph pengujian <i>Forecast</i> .....	49

## DAFTAR RUMUS

Rumus 2.1 Model <i>AR</i> .....	16
Rumus 2.2 Model <i>MA</i> .....	17
Rumus 2.3 Proses <i>ARMA</i> .....	18
Rumus 2.4 Proses <i>ARMA</i> .....	18
Rumus 2.5 Uji Stasioner .....	20
Rumus 2.6 Uji Stasioner .....	20
Rumus 2.7 Uji <i>ACF</i> .....	21
Rumus 2.8 Uji <i>PACF</i> .....	22

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang Masalah**

Transportasi dikatakan sebuah sarana yang menawarkan jasa perjalanan bagi setiap penggunanya. Seiring dengan semakin berkembangnya dunia transportasi maka semakin bertambah pula minat masyarakat menggunakan jasa transportasi terutama transportasi udara. Saat ini dunia transportasi udara adalah salah satu dari transportasi yang sangat pesat perkembangannya dibandingkan dengan transportasi lainnya. Tidak dapat dipungkiri penggunaan transportasi udara dikalangan masyarakat tidaklah asing lagi, semua itu dikarenakan murah nya harga yang ditawarkan pihak maskapai dan semakin tinggi nya tingkat kebutuhan masyarakat terhadap transportasi tersebut.

Setiap tahun nya jumlah masyarakat yang memakai jasa transportasi udara mengalami perkembangan yang semakin pesat. Hal tersebut kita ketahui bersama dengan semakin banyak nya maskapai penerbangan yang berlomba-lomba melakukan penawaran yang menarik perhatian masyarakat baik dari segi jangkauan harga yang ditawarkan maupun pelayanan dan fasilitas yang diberikan karena semakin ketatnya persaingan ditengah-tengah pihak maskapai penerbangan yang semakin marak pada saat ini.

Lion Air merupakan sebuah maskapai penerbangan dengan peminat yang cukup besar, hal tersebut dapat dilihat dari data yang telah didapatkan bahwa jumlah penumpang maskapai Lion Air lebih banyak dari maskapai lainnya, namun tidak dapat dipungkiri Lion Air juga sering mengalami penurunan penjualan tiket dengan berbagai faktor yang mendorong masyarakat tidak menggunakan maskapai tersebut sehingga jumlah penumpang mengalami ketidakstabilan di setiap waktu.

Suatu perusahaan didirikan dengan tujuan menghasilkan barang dan jasa sebagai suatu kebutuhan bagi konsumen sekaligus mendapatkan keuntungan dari usaha tersebut. Hal ini juga dipercaya dapat membantu pemerintah dalam meminimalisir tingginya angka tingkat pengangguran dengan membuka lapangan pekerjaan.

PT. Great Seasons Travel merupakan sebuah perusahaan yang bergerak dalam bidang biro jasa penjualan tiket pesawat. Perusahaan ini menjual tiket pesawat dari berbagai maskapai penerbangan. Perusahaan ini memiliki potensi dalam menyerap pembeli yang tidak stabil dari tahun ketahun nya. Hal tersebut dapat kita lihat dari ketidakstabilan jumlah penjualan yang hampir terus dialami.

Sehubungan dalam hal ini diperlukan untuk memprediksi seberapa besar peningkatan penjualan tiket yang didapatkan guna untuk memprediksi jumlah penjualan berupa ramalan ataupun perkiraan. Prediksi dilakukan hampir oleh semua pihak, baik pengusaha, pemerintah maupun masyarakat awam. Masalah yang diperkirakanpun bervariasi jenisnya, ada yang memperkirakan kemungkinan siapa pemenang dalam pemilihan presiden, perkiraan jumlah skor dalam sebuah

pertandingan olahraga, perkiraan cuaca, perkiraan jumlah arus lalu lintas, jumlah tingkat penjualan dan banyak lagi yang lainnya.

Metode *Autoregressive Integrated Moving Average* (ARIMA) atau yang disebut dengan metode runtutan waktu yang teknik dalam mencari sebuah pola yang paling sesuai dalam beberapa kumpulan data, dengan itu ARIMA sangat memanfaatkan sepenuhnya data dimasa lalu dan data dimasa sekarang dari variabel dependen dalam melakukan peramalan jangka pendek yang lebih akurat, sedangkan untuk peramalan jangka panjang ketepatan peramalannya tidaklah begitu baik, karena itu perlu dimengerti metode ARIMA adalah salah satu metode yang sangat baik ketepatannya dalam sebuah peramalan dikarenakan metode ini merupakan gabungan antara beberapa metode, metode ini juga hanya dapat menggunakan satu variabel sebagai salah satu dasar dalam melakukan sebuah prediksi.

Metode ARIMA dipilih sebagai metode yang digunakan untuk memprediksi dan mengkaji analisa jumlah penjualan tiket di PT.Great Seasons Travel, dengan menggunakan analisa prediksi ini diharapkan dapat mendukung proses penjualan tiket dikemudian hari. Berdasarkan keterangan diatas peneliti tertarik untuk melaksanakan sebuah penelitian yang berjudul **“DATA MINING MEMPREDIKSI JUMLAH PENJUALAN TIKET DENGAN MENGGUNAKAN METODE AUTOREGRESSIVE INTEGRATED MOVING AVERAGE “**

## **1.2 Identifikasi Masalah**

Identifikasi masalah yang diuraikan dalam penelitian ini adalah sebagai berikut :

1. Tidak stabil nya jumlah penjualan tiket yang terjadi pada PT.Great Seasons.
2. Semakin tinggi nya tingkat persaingan penjualan pada masa sekarang ini.

### **1.3 Batasan Masalah**

Berdasarkan penelitian tersebut maka batasan masalah pada penelitian ini adalah sebagai berikut :

1. Data yang digunakan adalah data penjualan dari bulan Mei 2018 sampai Oktober 2018.
2. Hanya memprediksi peningkatan penjualan tiket pada tahun 2019 dengan menggunakan aplikasi EViews 9
3. Hanya memprediksi data penjualan tiket dengan menggunakan maskapai Lion Air.

### **1.4 Rumusan Masalah**

Berdasarkan batasan masalah yang telah dikemukakan sebelumnya maka rumusan masalah yang dapat diperoleh dari penelitian ini adalah sebagai berikut :

1. Bagaimana mengolah data penjualan tiket pada PT.Great Seasons Travel?
2. Bagaimana penerapan metode ARIMA dapat memberikan model terbaik untuk memprediksi jumlah peningkatan penjualan tiket?
3. Bagaimana mendapatkan prediksi dengan mengaplikasikan tools EViews 9?

### **1.5 Tujuan Penelitian**

Adapun tujuan yang ingin dicapai dari penelitian ini adalah sebagai berikut:

1. Untuk Memprediksi jumlah peningkatan penjualan tiket pesawat dengan melihat data – data sebelumnya.
2. Untuk mengetahui implementasi metode ARIMA dalam memprediksi peningkatan penjualan tiket pesawat.
3. Mengaplikasikan Eviews 9 untuk mendapatkan prediksi dengan menggunakan metode ARIMA.

### **1.6 Manfaat Penelitian**

Manfaat penelitian dapat dikategorikan menjadi 2 yaitu :

1. Aspek Teoritis
  - a. Mengetahui langkah apa yang akan dilakukan selanjutnya untuk lebih meningkatkan jumlah penjualan tiket dimasa yang akan datang.
  - b. Mengetahui apakah maskapai penerbangan Lion Air mengalami peningkatan maupun penurunan.

2. Aspek Praktis

- a. Bagi Perusahaan

Mengetahui secara detail prediksi peningkatan penjualan tiket pesawat.

- b. Bagi Kampus

Sebagai sumber referensi atau informasi tambahan tentang metode ARIMA untuk siswa yang ingin melakukan penelitian lebih lanjut.

- c. Peneliti

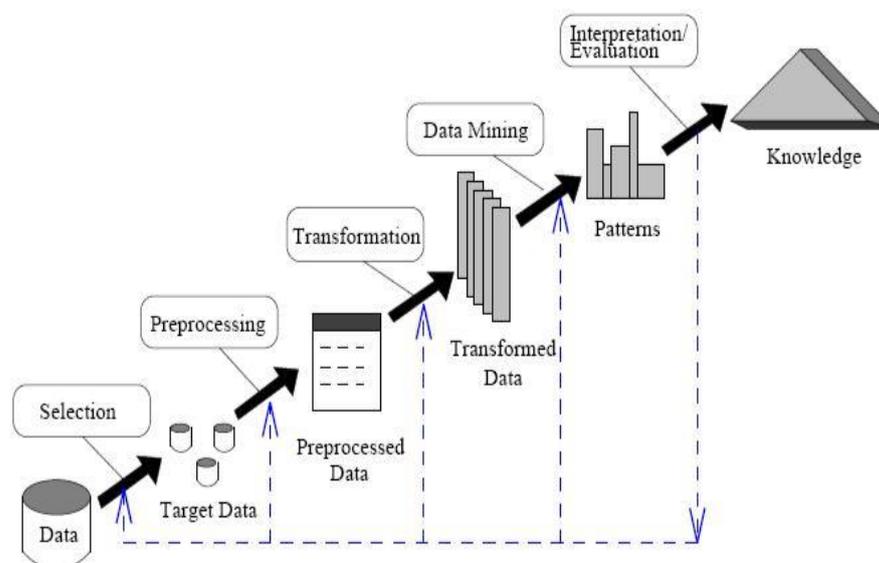
Untuk mengetahui dan meningkatkan pengetahuan tentang Penambangan Data menggunakan metode rata-rata bergerak terintegrasi *Autoregressive* dan juga sebagai salah satu persyaratan utama dalam memperoleh gelar Sarjana Teknik Informatika.

## BAB II

### KAJIAN PUSTAKA

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

KDD merupakan suatu kegiatan dimana dalam hal ini meliputi kegiatan pemakaian historis data yang bertujuan menemukan historis data, pengumpulan data, menentukan pola dan hubungan dalam suatu data yang besar. KDD sering sekali digunakan dalam menjabarkan proses pencarian sebuah informasi dalam sebuah database yang besar. Pada saat ini KDD semakin berkembang, karena pengembangan itu berdampak mengakibatkan penggunaan *pattern recognition* telah berkurang dikarenakan KDD sudah merupakan bagian dalam *Data Mining*. Hasil dari *Data Mining* tersebut dapat dipakai dalam mengambil sebuah keputusan dimasa yang akan datang.



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

Tahapan penelitian dengan menggunakan proses *knowledge discovery in database ( KDD )* adalah sebagai berikut :

a. *Data Selection* (Seleksi Data)

Pada tahapan ini penyeleksian data bersumber dari data yang diperoleh dari data penjualan selama 6 bulan pada Pt Great Seasons, data yang akan diambil mengandung kategori yang serupa dengan referensi yang telah ditentukan. Pada tahapan ini juga data yang telah dikelompokkan sesuai kategori masing-masing akan difokuskan pada subset variabel atau sampel data dimana data tersebut akan menjadi data operasional yang akan digunakan dalam proses data.

b. *Preprocessing* (Pembersihan data)

Tujuan tahapan ini adalah untuk memperoleh hasil yang akan digunakan sebagai titik fokus dalam penelitian, membuang seluruh penggandaan data yang ada, memeriksa isi data serta memperbaiki secara keseluruhan kesalahan yang ada pada data .

c. *Transformation* (Transformasi Data)

Pada tahapan transformasi ini data yang berada dalam *Data Mining* sangatlah perlu dilakukan yang namanya transformasi data dikarenakan dalam proses *Data Mining* data yang belum valid akan diintegrasikan, dimana peranannya adalah membuat data yang belum memiliki entitas yang jelas akan diubah menjadi sebuah data yang valid karena data yang diperlukan dalam proses *Data Mining* adalah data yang valid.

d. *Data Mining*

Pada tahapan ini data dipusatkan demi mendapatkan sebuah pola dari data penjualan yang sebelumnya telah ditransformasikan dengan menggunakan algoritma yang akan dipakai.

e. *Interpretation* (evaluasi)

Pada tahap ini, KDD melakukan penyaringan untuk mengetahui apakah informasi atau data itu dapat berbanding terbalik dengan fakta yang telah dijelaskan sebelumnya. Informasi atau pola yang ditemukan akan ditampilkan dalam bentuk yang mudah dipahami dan dipahami oleh pengguna.

## **2.2 Data Mining**

### **2.2.1 Definisi Data Mining**

*Data mining* adalah sebuah proses dimana proses yang dilakukan untuk menentukan sebuah pola serta informasi dari sekelompok data yang berjumlah besar ( Pramudiono,2006 ). *Data Mining* di desain demi dapat memadai sebuah basis data yang bertujuan untuk memperoleh informasi yang dapat digunakan yang sebelumnya tidak diketahui kegunaannya. *Data Mining* didefinisikan sebagai tahap yang digunakan untuk mencari dan menemukan hubungan antara pola dan tren baru yang berguna untuk memfilter data skala besar yang disimpan dalam database. Penambangan Data juga dapat disebut sebagai mengekstraksi informasi baru yang diambil dari sejumlah besar data yang membantu dalam membuat keputusan. Istilah *Data Mining* terkadang juga disebut *Knowledge discovery*.

### 2.2.2 Pengelompokan *Data Mining*

Pengelompokan Data mining terdiri beberapa kelompok berdasarkan tugas pemrosesan yang dapat dilakukan yaitu :

1. Deskripsi

Deskripsi ialah cara pengaplikasian pola dan kecenderungan yang dimiliki pada suatu data. Dalam hal ini seringkali peneliti memiliki kecenderungan hanya memberikan kemungkinan penjelasan dari suatu pola dikarenakan sulitnya menentukan keterangan atau fakta mengenai data yang diperoleh.

2. Estimasi

Estimasi merupakan sebuah model yang dirancang dengan memakai *record* data lengkap yang telah tersedia nilai variabel targetnya sebagai nilai prediksi yang cenderung lebih mengarah ke numerik dibandingkan ke arah kategori.

3. Prediksi

Prediksi merupakan suatu keadaan dimana nilai sebuah data dihasilkan dengan menerka sebuah nilai yang belum diketahui kepastian nya dengan menggunakan rentet waktu guna memperkirakan nilai dimasa depan.

4. Klasifikasi

Klasifikasi memiliki sebuah kategori target variabel yang berfungsi sebagai pemisah atau sebagai penentu golongan pada tiap kategori.

5. *Pengklusteran*

*Pengklusteran* yaitu sebuah pengelompokan *record* maupun pengamatan yang dapat berbentuk sebuah kelas objek-objek yang memiliki tingkat kemiripan yang hampir sama dimana dalam hal ini pengklusteran tidak memiliki variabel target dikarenakan pengklusteran tidak melakukan estimasi atau prediksi maupun klasifikasi terhadap data yang diperoleh.

#### 6. Asosiasi

Fungsi asosiasi yang mengambil atribut yang ditampilkan sekaligus. Analisis asosiasi ini juga dikenal sebagai salah satu teknik penambangan data yang sangat berfungsi sebagai dasar untuk berbagai teknik penambangan data yang saat ini ada.

### 2.3 Prediksi

Peramalan ialah seni dan ilmu untuk memprediksi kejadian dimasa mendatang. Prediksi (*forecasting*) merupakan suatu alat atau teknik yang digunakan untuk memprediksi sebuah nilai pada masa yang akan datang dengan memperhatikan data dan informasi yang relevan, baik data maupun data informasi masa lalu atau data pada masa sekarang. Terdapat 2 metode dalam prediksi, diantaranya :

#### 1. Metode Kualitatif

Metode ini adalah metode yang digunakan di mana tidak ada model matematika, ini karena data yang diperoleh tidak cukup representatif untuk memprediksi masa depan (peramalan jangka panjang). Peramalan *kualitatif* biasanya menggunakan pendapat ahli dalam bidangnya.

Metode ini juga memiliki kelebihan dimana metode ini tidak memerlukan banyak biaya karena tidak membutuhkan data serta cepat diperoleh. Sementara kekurangan dari metode tersebut yaitu data yang dihasilkan bersifat subjektif.

## 2. Metode Kuantitatif

Metode ini adalah metode yang didasari dengan penggunaan data mentah yang disertai dengan kaidah matematis untuk menghasilkan ramalan dimasa depan. Terdapat beberapa macam model peramalan dalam metode kuantitatif tersebut, yaitu :

### a. Model-model regresi

Model ini merupakan perpanjangan dari model regresi linier di mana model ini memprediksi sebuah variabel yang terdapat hubungan linier dengan variabel independen yang diketahui dan dapat diandalkan.

### b. Model Ekonometrik

Model ini menggunakan serangkaian persamaan regresi di mana ada variabel independen yang merangsang segmen ekonomi

### c. Model *Time Series Analysis* (Deret Waktu)

Ini adalah model di mana model ini menempatkan garis tren yang *representatif* dengan data dimasa lampau berdasarkan tren data dan mengimplementasikan data ke masa depan.

#### **2.3.1 *Time Series Analysis***

Data deret waktu adalah data yang merupakan kumpulan data dalam periode waktu tertentu. Peramalan deret waktu juga merupakan perkiraan berdasarkan perilaku data masa lalu yang diimplementasikan ke masa depan dengan

menggunakan persamaan matematika dan statistik. Tipe data deret waktu dibagi menjadi beberapa jenis, termasuk (Arsyad, 2001):

a. Siklus

Pola ini adalah serangkaian perubahan naik dan turun sehingga pola siklus ini berubah dan bervariasi. Pola siklus dan pola ireguler diperoleh dengan menghilangkan tren atau pola musiman jika data yang digunakan adalah dalam bentuk mingguan atau bulanan.

b. *Random*

Pola acak yang tidak bisa digambar. Pola ini disebabkan oleh peristiwa yang tidak terduga sehingga polanya tidak teratur dan tidak dapat diprediksi yang bersifat acak sehingga tidak dapat dijelaskan.

c. *Trend*

Tren adalah kecenderungan komponen jangka panjang tertentu dalam pola data, baik dalam hal arah kenaikan dan penurunan dari waktu ke waktu sehingga pola tren tidak konstan.

d. Musiman

Pola yang menunjukkan gerakan yang diulang secara teratur dari satu periode ke periode berikutnya. Pola ini dapat ditunjukkan oleh data plak yang telah diklasifikasikan baik mingguan dan bulanan. Teknik yang biasa digunakan dalam menentukan nilai-nilai pola musiman adalah metode *moving averages*, *smoothing exponential* musim dingin klasik.

Teknik peramalan *Time series* terdiri atas:

1. Statistika

- a. *Moving Average*
- b. *Exponential Smoothing*
- c. Regresi
- d. *ARIMA ( Box Jenkins )*
- 2. Kecerdasan buatan
  - a. *Neural Network*
  - b. Algoritma Genetika
  - c. *Simulated Annealing*
  - d. *Genetic Programming*
  - e. Klasifikasi
  - f. *Hybrid*

## **2.4 Pengertian Penjualan**

Salah satu kegiatan yang dilakukan oleh manusia saat ini untuk memenuhi berbagai kebutuhan dan keinginan mereka adalah berjualan. Menurut T Rhonda Abrams “Penjualan adalah kegiatan yang berinteraksi langsung dengan konsumen untuk mendapatkan pesanan atau penjualan langsung”. Penjualan juga memiliki fungsi pertukaran dimana penjualan merupakan salah satu aktivitas manusia yang bergerak di bidang bisnis dan ekonomi demi mendapatkan suatu keuntungan. Maskapai Lion Air juga melakukan proses penjualan dimana masyarakat dapat menggunakan maskapai tersebut dengan membeli tiket yang dikeluarkan oleh maskapai tersebut di berbagai agen maupun dari pihak maskapai langsung, begitu juga dengan PT.Great Seasons yang merupakan sebuah perusahaan agen perjalanan yang berfungsi sebagai media untuk maskapai tersebut melakukan penjualan.

### **2.4.1 Faktor Pengaruh Penjualan**

Dalam melakukan transaksi penjualan terdapat beberapa faktor yang mempengaruhi antar kedua pihak :

#### **1. Kemampuan Penjualan**

Dalam hal ini kemampuan dari si penjual merupakan salah satu hal yang penting dikarenakan hal tersebut dapat memberi keyakinan kepada pembeli untuk membeli barang mereka

#### **2. Kondisi Pasar**

Kondisi pasar juga merupakan masalah penting. Beberapa hal yang harus diperhatikan dalam mengamati kondisi pasar adalah kelompok pembeli, segmentasi pasar, frekuensi pembelian, keinginan, dan daya beli.

#### **3. Modal**

Modal merupakan biaya pertama yang dikeluarkan si penjual dalam penyediaan produk dan mendapatkan barang-barang yang akan ditawarkan untuk seluruh konsumen.

### **2.5 Metode *Autoregressive Integrated Moving Average*(ARIMA)**

Metode *Autoregressive Integrated Moving Average* (ARIMA) adalah sebuah metode yang disebut dengan metode runtun waktu dimana metode ini adalah model metode yang diterapkan untuk analisa deret , peramalan dan pengendalian dengan mengabaikan secara penuh variabel independent dalam menetapkan sebuah keputusan peramalan. Metode ARIMA juga merupakan metode yang menggunakan nilai lampau dan saat ini dari variabel dependen dalam menghasilkan estimasi

jangka pendek yang akurat. ARIMA juga cocok jika pengamatan dari deret waktu secara statistika terkait satu sama lain (tergantung). Metode ini merupakan salah satu model yang sangat populer dalam peramalan dengan menggunakan sebuah pendekatan yang disebut dengan pendekatan *time series*.

### 2.5.1 Klasifikasi Model ARIMA

Model Box-Jenkins (ARIMA) terbagi menjadi 3 kategori: Model *Autoregressive*(AR), *Moving Average*(MA), dan model campuran ARIMA yang merupakan kombinasi karakteristik dari kedua model sebelumnya..

#### 1. *Autoregressive Model* (AR)

*Autoregressive* adalah bentuk di mana model ini menyatakan prediksi yang digunakan menjadi fungsi dari nilai sebelumnya dari deret waktu tertentu (Makridakis, 1995: 513), model ini juga model yang dalam bentuk regresi nya dapat menghubungkan nilai-nilai lainnya dengan dirinya sendiri. Bentuk umum model ini dengan ordo p (AR(p)) atau model ARIMA (p,0,0). Bentuk umum model AR (p) adalah:

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \epsilon_t$$

**Rumus 2.1 Model AR**

dimana:

$X_t$  : nilai variabel pada waktu ke-t

$X_{t-1}, X_{t-2}, \dots, X_{t-p}$  : nilai dari *time series* yang bersangkutan pada waktu t, t-1, t-2, ..., t-p

$\phi_i$  : koefisien regresi, i: 1, 2, 3, ..., p

$\varepsilon_t$  : nilai *error* pada waktu ke- $t$

$p$  : Orde AR

## 2. Moving Average Model(MA)

Bentuk dari model ini ialah orde  $q$  (MA( $q$ )) atau *ARIMA* (0,0, $q$ ) dinyatakan sebagai berikut:

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}; \varepsilon_t \sim N(0, \sigma^2)$$

### Rumus 2.2 Model MA

dimana:

$X_t$  : jumlah variabel pada waktu ke- $t$

$\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$  : nilai-nilai dari *error* pada waktu  $t, t-1, t-2, \dots, t-q$

$\theta_i$  : koefisien regresi ,  $i: 1, 2, 3, \dots, q$

$q$  : orde *MA*

## 3. Model Campuran

### a. Proses ARMA

sebuah proses dimana bentuk umum yang digunakan untuk campuran proses AR(1) murni dan MA(1) murni dan bentuk umum dari proses AR dan MA ( $p,q$ ), dapat dinyatakan sebagai berikut:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

**Rumus 2.3** Proses ARMA

- $X_t$  : nilai variabel pada waktu ke-t
- $\theta_i$  : parameter model MA ke-i, i: 1, 2, 3, ..., q
- p : Orde AR
- q : orde MA
- $\phi_i$  : koefisien regresi, i: 1, 2, 3, ..., p
- $\varepsilon_t$  : nilai *error* pada waktu ke-t
- $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$  : Error pada waktu  $t, t-1, t-2, \dots, t-q$  dan  $\varepsilon_t$  diasumsikan sebagai *white noise* dan normal.

b. Model ARIMA

Merupakan sebuah proses dimana jika non stasioneritasnya ditambahkan pada campuran proses ARMA, maka model umum dari proses ARIMA ( $p, d, q$ ) dapat dipenuhi. Bentuk umum dari proses ini dapat dinyatakan sebagai berikut :

$$i\phi B (1 - B)^d X_t = \theta(B) \varepsilon_t; \varepsilon_t \sim N(0, \sigma^2) \quad \text{Rumus 2.4 Proses ARMA}$$

Persamaan diatas tersebut dapat ditulis dengan menggunakan operator B (Backshift).

$$(1 - B)^d (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \varepsilon_t$$

Sehingga diperoleh :

$$(1 - B)^d (X_t - \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p}) = (\epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q})$$

- $X_t$  : nilai variabel pada waktu ke-t
- $(1 - B)^d X_t$  : *time series* yang stasioner pada pembedaan ke-d
- B : Operator *backshift*
- p : Orde AR
- q : orde MA
- d : orde pembedaan
- $\epsilon_t$  : nilai *error* pada waktu ke-t
- $\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-q}$  : *Error* pada waktu  $t, t-1, t-2, \dots, t-q$  dan  $\epsilon_t$  diasumsikan sebagai *white noise* dan normal

Asumsi stasioneritas dapat diartikan sebagai sebuah perubahan pada sebuah data, jika data tersebut mengalami stasioneritas maka terdapat suatu perubahan yang terjadi secara drastis pada suatu data. Rangkaian waktu dapat disebut stasioner jika rata-rata dan varians dalam data konstan, tidak ada elemen tren dan elemen perubahan musim. Musiman dapat diartikan sebagai suatu kondisi di mana data memiliki pola yang diulang dalam periode waktu tertentu. Dalam data stasioner faktor yang sering menentukan data musiman adalah koefisien autokorelasi pada dua atau tiga jeda waktu yang secara signifikan berbeda dari nol.

### 2.5.2 Mengidentifikasi Model

*AR* dan *MA* dari model arima hanya berhubungan dengan seri periodik stasioner. Stasioner berarti tidak ada pertumbuhan atau penurunan data. Suatu data yang tidak stasioner harus diubah menjadi stasioner dengan membuat perbedaan yang dilakukan untuk menghitung selisih nilai observasi.

$$X'_t = X_t - X_{t-1}$$

**Rumus 2.5** Uji Stasioner

Keterangan :

$X'_t$  : nilai deret berkala setelah differencing

$X_t$  : nilai deret berkala pada waktu  $t$

Persamaan diatas dapat ditulis kembali menjadi :

$$X'_t = X_t - BX_{t-1}$$

**Rumus 2.6** Uji Stasioner

Serangkaian berskala dikatakan diam jika menghasikan kesalahan acak jika koefisien autokorelasi untuk semua kelambatan adalah angka yang ditunjukkan pada setiap interval yang secara statistik tidak berbeda dari nol atau berbeda dari nol tetapi untuk beberapa kelambatan di depan. Koefisien autokorelasi dikatakan tidak berbeda dari nol jika berada dalam interval.

$$0 \pm Z_{\alpha/2} (1 / \sqrt{n})$$

Keterangan :

$Z_{\alpha/2}$  : Nilai variabel normal standart dengan  $1 - \alpha$

$N$  : Banyak data

Setelah data yang diperoleh stasioner maka selanjutnya model ARIMA ditentukan dengan melihat autokorelasi dan partial autokorelasi

### 1. Autocorelation Function (ACF)

Merupakan korelasi antara data periode waktu  $t$  dengan periode waktu  $t-1$  sebelumnya .

$$r_k = \frac{\sum_{t=b}^{n-k} (X_t - X_{rata})(X_{t+k} - X_{rata})}{(X_t - X_{rata})^2} \quad \text{Rumus 2.7 Uji ACF}$$

$r_k$  : Nilai sampel ACF pada lag  $k$

$X_t$  : Nilai deret berkala pada waktu  $t$

$X_{rata}$  : rata-rata dari data berkala

Nilai standart error ACF pada Lag  $k$  dapat dilihat dibawah ini :

$$S_{rk} = \frac{(1 + 2 \sum rk^2)^{1/2}}{(n-b+1)^2}$$

Keterangan :

$S_{rk}$  : Nilai standart error pada lag  $k$

$b$  : awal periode dari data

Nilai ACF pada lag dapat dilihat :

$$T_{rk} = \frac{rk}{S_{rk}}$$

### 2. Partial Autocorelaton Function (PACF)

Partial dipakai dalam menghitung tingkat kecerdasan antara  $X_t$  dan  $X_{t-k}$  apabila *lag time* dianggap terpisah.

$$r_{kk} = \frac{r_k - \sum r_{k-1}}{1 - \sum r_{k-1}} \quad \text{Rumus 2.8 Uji PACF}$$

$r_{kk}$  : Nilai sampel PACF pada ordo ke k

$r_k$  : Nilai sampel ACF pada lag ke k

Nilai standart error PACF berorde k seperti berikut :

$$S_{r_{kk}} = \frac{1}{(n-b+1)^{1/2}}$$

Keterangan :

$S_{r_{kk}}$  : Nilai standart error PACF pada lag k

$n$  : jumlah data

$b$  : awal periode data

Nilai PACF berorde k

$$t_{r_{kk}} = \frac{r_{kk}}{S_{r_{kk}}}$$

keterangan :

$t_{r_{kk}}$  : Nilai statistic PACF pada ordo ke k

$r_{kk}$  : Nilai sampel PACF pada ordo ke k

$S_{r_{kk}}$  : Nilai standart error PACF pada lag k

## 2.6 Software Pendukung



**Gambar 2.2** Logo Eviews 9

EViews adalah aplikasi yang digunakan dalam pendidikan, pemerintahan dan perindustrian. EViews atau *Views of Econometrics*, adalah versi terbaru dari paket statistik untuk memanipulasi data deret waktu. Meskipun EViews dirumuskan oleh para ekonom, program ini juga dapat digunakan dalam bidang studi, seperti sosiologi, statistik, keuangan. Perangkat lunak ini dikembangkan oleh perusahaan Quantitative Micro Software (QMS) pada tahun 1994. Aplikasi ini menggunakan lingkungan windows yang ramah pengguna, banyak perusahaan yang beroperasi dilakukan dengan menu drop-down. Penggunaannya terdiri dari analisis dan evaluasi data, analisa keuangan, prakiraan ekonomi makro, simulasi, prakiraan penjualan, dan analisis biaya..

## 2.7 Penelitian Terdahulu

1. **Penerapan Metode ARIMA untuk Peramalan Pengunjung Perpustakaan UIN SUSKA RIAU** (Ahmad Sayuti, Mustakim, Syarfi Aziz, 2017)

Setiap tahun jumlah mahasiswa UIN Suska Riau mengalami peningkatan, hal ini juga mempengaruhi jumlah pengunjung ke Perpustakaan Universitas. Meningkatnya Jumlah pengunjung perpustakaan harus disertai dengan pengembangan sarana dan prasarana yang memadai. Oleh karena itu, peramalan diperlukan untuk mendukung perencanaan dan pengembangan Perpustakaan. Studi ini membahas tren perkembangan jumlah pengunjung ke Perpustakaan UIN Suska Riau di masa mendatang. Data yang digunakan dalam penelitian ini adalah data pengunjung perpustakaan dari Januari 2016 hingga Agustus 2016. Analisis data menggunakan metode ARIMA. Metode ini melakukan identifikasi dalam menentukan estimasi parameter, menentukan model terbaik dengan uji statistik dan prediksi untuk data di masa depan. Alat yang digunakan dalam penelitian ini adalah Eviews. Hasil analisis dari penelitian ini adalah penggunaan model ARIMA untuk data pengunjung di Perpustakaan UIN SUSKA Riau, yang merupakan model ARIMA (0,1,1) dan diperoleh hasil perkiraan hingga hari ke-190, yaitu Departemen Informatika Teknik 10 Siswa, Teknik Industri 1 Siswa, Sistem Informasi 3 Siswa, Matematika Terapan 3 Siswa, Teknik Elektro 2 Siswa dan total 24 Siswa per hari. Dari hasil penelitian yang telah dilakukan, dapat ditarik kesimpulan bahwa metode ARIMA dapat digunakan untuk memprediksi pengunjung perpustakaan UIN SUSKA Riau. Model MA dalam dilakukan dalam penelitian ini lebih sesuai untuk data pengunjung perpustakaan UIN SUSKA Riau. Hasil dari penelitian ini adalah prediksi pengunjung perpustakaan UIN Suska Riau dihari ke 190 yaitu TIF 10, TIN 1, SIF 3, MATER 3, TE 2, dan total perhari 24 orang.

## 2. **Prediksi Saham Dengan Menggunakan Metode Autoregressive dan Algoritma Kelelawar** (Bayu Prabawa, Jondri Nasri, Mahmud Dwi Sulisty, 2015)

Tabungan merupakan suatu faktor yang konvensional. Para pemilik saham telah menemukan lahan investasi baru yang lebih menjanjikan dari pada deposito ataupun tabungan di bank. Lahan investasi baru tersebut adalah investasi di bursa saham. Namun, harga saham berfluktuasi setiap hari, oleh karena itu diperlukan suatu sistem yang dapat memberikan prediksi untuk harga saham ini untuk membantu investor mengambil tindakan yang tepat sehingga risiko dapat diminimalkan.. Metode yang digunakan yaitu metode autoregressive untuk model prediksinya dan algoritma kelelawar untuk optimasinya. Metode autoregressive diperkenalkan oleh Box dan Jenkins, metode yang sering digunakan untuk pemodelan data time series dalam peramalan. Untuk algoritma kelelawar, diperkenalkan oleh Xin-She Yang. Algoritma ini adalah salah satu algoritma metaheuristik yang terinspirasi oleh perilaku echolocation dari kelelawar. Algoritma ini dikembangkan berdasarkan kelebihan-kelebihan dari algoritma PSO, GA, SA dan pencarian harmonys sehingga kesimpulan yang didapat adalah Metode *autoregressive* dan algoritma kelelawar dapat menyelesaikan masalah prediksi harga saham dengan tingkat error dibawah 4% kecuali pada pengujian dengan menggunakan 10 data untuk prediksi saham amazon dan facebook bulanan tingkat error melebihi 4% sehingga Sistem ini masih merupakan model dan belum dapat diimplementasikan pada investasi saham yang sebenarnya, karena masih memiliki nilai MAPE yang cukup tinggi.

### 3. Penerapan *Data Mining* untuk Analisa Pola Pembelian Konsumen Dengan Algoritma *FP-Growth* pada Data Transaksi Penjualan Spare Part Motor

(Alfannisa Annurullah Fajrin, AlgiFanri Maulana, 2018)

Di perusahaan yang memiliki banyak cabang atau dealer seperti CV.TJAHAJA BARU, membutuhkan lokasi yang harus dipertimbangkan dalam mendirikan cabang baru, harus dapat mempengaruhi pola pembelian konsumen, karena pola pembelian masing-masing konsumen berbeda. Ini perlu dianalisis lebih lanjut sehingga dapat menghasilkan informasi yang bermanfaat, dan memaksimalkan manfaat yang bisa diperoleh. Penambangan Data dapat digunakan oleh perusahaan besar untuk menggali data untuk mendapatkan informasi yang dapat mendukung dan meningkatkan proses bisnis perusahaan. Jadi dalam penelitian ini, pengujian dilakukan dengan menggunakan algoritma *FP-Growth* untuk membantu perusahaan mengetahui pola pembelian konsumen dan transaksi penjualan suku cadang, sehingga dapat disimpulkan bahwa penerapan metode *Data Mining* dengan algoritma *FP-Growth* ke dalam aplikasi untuk analisis pola pembelian konsumen sangat bermanfaat bagi perusahaan, karena TB-Damar akan mengetahui suku cadang mana yang sedang dibeli dan membantu dalam memesan suku cadang di kantor pusat. Hasil penjualan sparepart motor sport paling banyak dijual di TB-Damar CV. TJAHAJA BARU Padang dapat diketahui menggunakan algoritma *FP-Growth*. *Spare part* yang memenuhi *minimum support* dan *minimum confidence* serta yang banyak terjual adalah *screw valve adjusting*, *oil seal*, *battery assy*, *axle*, *gasket cylinder*, dan *cable clutch*.

#### **4. Metode *Data Mining* Untuk Seleksi Calon Mahasiswa Pada Penerimaan Mahasiswa Baru Di Universitas Pamulang (Aries Saifudin,2017)**

Universitas berusaha menawarkan pendidikan tinggi dengan harga yang terjangkau bagi kelas bawah. Namun, banyak siswa yang drop out setiap semester, menghasilkan rasio yang berbeda antara jumlah siswa baru dan jumlah lulusan yang tidak seimbang. Selain itu, banyak siswa tidak lulus tepat waktu, yang menyebabkan rasio yang tidak sama antara dosen dan mahasiswa. Keduanya mengurangi peringkat pada saat akreditasi. Alasan untuk menghentikan siswa yang belum menyelesaikan pendidikannya atau tidak dapat menyelesaikan pelatihan tepat waktu tidak dapat ditentukan dengan sistem seleksi saat ini. Studi ini menyarankan penggunaan teknik-teknik penambangan data untuk memprediksi tingkat kemahiran siswa potensial. Dengan menggunakan teknik penambangan data dan pembelajaran mesin, prediksi dapat dibuat berdasarkan data masa lalu. Metode Data Mining yang digunakan untuk memprediksi adalah klasifikasi, yaitu Naïve Bayes (NB), k-Nearest Neighbor (k-NN), Hutan Acak (RF), Decision Stump (DS), Decision Tree (DT), Induksi Aturan (RI), LineAR Regression (LR), Linear Discriminant Analysis (LDA), Neural Network (NN), dan Support Vector Machine (SVM). Berdasarkan hasil implementasi dan pengukuran algoritma / model yang diusulkan, diperoleh algoritma / model terbaik, yaitu Support Vector Machine (SVM) dengan akurasi 65,00%. Tetapi akurasi ini masih jauh dari sempurna.

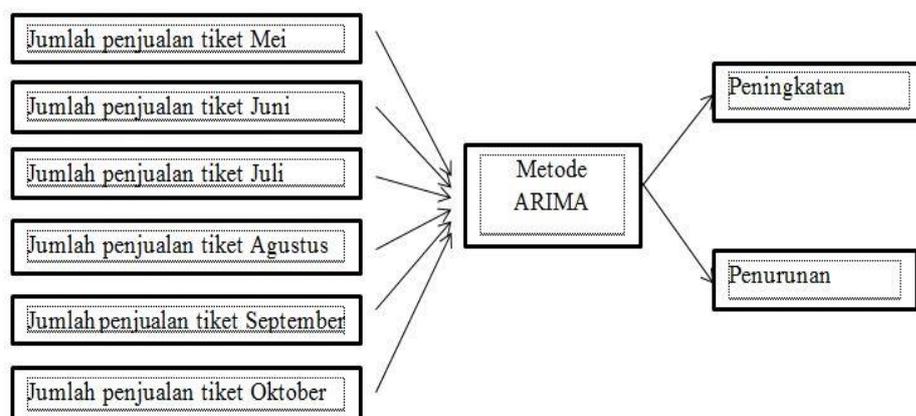
**5. Analisis Tingkat Penerimaan Calon Konsumen Terhadap Jenis Mobil Dengan Menggunakan Metode Regresi Linier** (Nur Intan Permata Hati, Yusuf Sulistyono Nugroho, 2016).

Di zaman modern, di mana orang saat ini memiliki berbagai peluang kerja, orang membutuhkan alat transportasi untuk mendukung peluang kerja mereka sehingga mereka dapat berjalan dengan baik. Salah satu alat transportasi yang paling umum adalah mobil. Jumlah penggemar mobil meningkatkan permintaan untuk jenis transportasi ini. Selain itu, peningkatan jumlah mobil penumpang di Indonesia juga disebabkan oleh faktor peningkatan pendapatan dari kelas menengah ke atas. Hal ini mengakibatkan banyak pembuat mobil di Indonesia berlomba menawarkan mobil dengan model yang berbeda dan, tentu saja, manfaat yang lebih menarik dengan harga yang semakin kompetitif. Terhadap latar belakang ini, penelitian ini menganalisis tingkat penerimaan konsumen potensial model mobil menggunakan teknik data mining menggunakan metode regresi linier. Metode ini bertujuan untuk memprediksi nilai fungsi kepadatan probabilitas dan turunan kedua dari suatu titik. Hasil yang diharapkan dari penelitian ini adalah dalam bentuk model regresi linier dalam bentuk formula yang dapat menentukan penerimaan konsumen tipe mobil potensial di pasar sehingga perusahaan mobil dapat memahami kebutuhan konsumen tentang mobil berdasarkan harga pembelian, harga pembelian (harga perawatan (harga perawatan) ), Jumlah pintu (jumlah pintu), kapasitas (kapasitas), ukuran bagasi (ukuran kompartemen bagasi) dan perkiraan keselamatan dengan variabel dependen Penerimaan Otomatis. Model regresi linier yang diproduksi pertama kali diuji sehingga model tersebut akan valid

untuk digunakan lebih lanjut sehingga berdasarkan penelitian yang dilakukan dapat disimpulkan bahwa dalam pengujian Korelasi dan uji t-statistik model regresi yang dihasilkan akan lemah. Tetapi jika uji determinasi, uji f-statistik, uji normalitas, uji multikolinieritas dan interpretasi hasil regresi adalah valid dan masih dapat digunakan.

## 2.2 Kerangka Pemikiran

Untuk memprediksi jumlah penjualan tiket pada PT. Great Seasons travel pada musim liburan perlu dilakukannya analisis data dengan menggunakan Metode ARIMA, dalam penelitian ini data-data sebelumnya diolah dengan menggunakan *software* Eviews guna mendapatkan hasil prediksi tingkat penjualan tiket yang lebih akurat. Dari latar belakang dan metode yang digunakan maka kerangka pemikiran dari penelitian adalah:



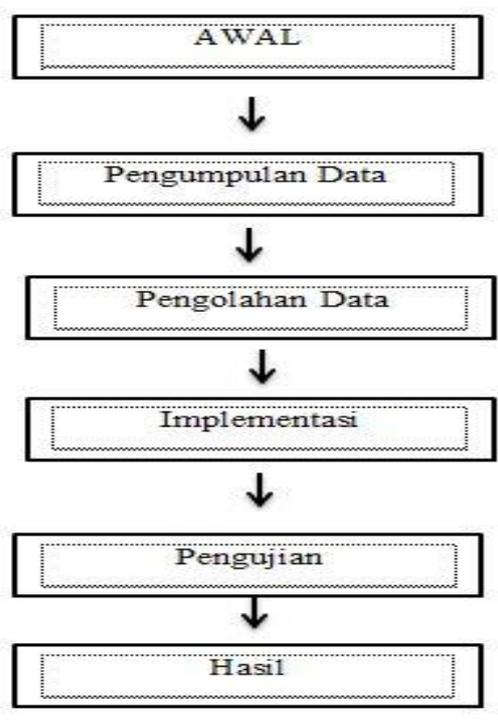
**Gambar 2.3** Kerangka pemikiran

## BAB III

### METODE PENELITIAN

#### 3.1 Desain Penelitian (Arsitektur Sistem)

Desain penelitian yang digunakan dalam penelitian tersebut merupakan gambaran langkah-langkah yang dilakukan mulai dari awal penelitian dilaksanakan sampai pada akhir penelitian tersebut dilaksanakan.



**Gambar 3.1** Desain Penelitian

Dari desain penelitian yang peneliti gambarkan diatas, dapat peneliti jelaskan sebagai berikut:

1. Awal

Pada tahap awal penelitian ini, memahami dan memahami dasar masalah yang akan digunakan sebagai studi. Tentukan ruang lingkup masalah, latar belakang masalah dan pelajari beberapa teori yang berkaitan dengan penelitian dan bagaimana menemukan solusi untuk masalah dan hasil dalam penelitian untuk mencapai tujuan yang diinginkan.

## 2. Pengumpulan Data

Pada tahap ini, data yang diolah dalam penelitian ini diperoleh dari laporan bulanan penjualan PT. Great Seasons dalam jangka waktu 6 bulan pada 2018. Data tersebut merupakan data mentah (*raw data*) yang akan digunakan sebagai data yang akan menghasilkan data yang dibutuhkan oleh *Data Mining* untuk diubah menjadi sebuah format yang dapat diproses *Data Mining* ketahap selanjutnya.

## 3. Pengolahan Data

Pengolahan data merupakan keseluruhan inputan data dalam bentuk yang berupa informasi, Kemudian data yang telah dipilih dan diseleksi selanjutnya diproses dengan Input dalam perancangan sistem penerapan *Data Mining* prediksi penjualan tiket dengan menggunakan metode autoregressive integrated moving average merupakan rekapitulasi transaksi lalu akan diuji kembali dengan menggunakan aplikasi Eviews 9.

## 4. Implementasi

Pada tahapan ini penulis menggunakan aplikasi Eviews 9 demi mempermudah penulis dalam membuktikan hasil analisa penelitian yang dilakukan. Hasil tersebut akan digunakan oleh PT.Great Seasons Travel demi peningkatan penjualan .

## 5. Pengujian

Merupakan tahapan dimana data-data yang telah dikumpulkan kemudian akan di uji dengan menggunakan aplikasi Eviews 9 dimana hasil yang didapatkan akan menjadi kesimpulan dari data yang di teliti. Proses pengujian dimulai dengan penginputan data yang kemudian diolah dengan aplikasi Eviews 9

## 6. Hasil

Tahap ini merupakan akhir dari penelitian dimana data yang telah diolah menghasilkan informasi yang memiliki nilai akurasi yang tinggi dari prediksi.

### **3.2 Teknik Pengumpulan Data**

Metode merupakan suatu cara atau teknik dalam pengumpulan data yang dilakukan secara sistematis dalam menyelesaikan suatu kasus. Untuk itu penulis menggunakan beberapa cara untuk memperolehnya, diantaranya yaitu :

#### 1. Studi Lapangan (Field Research)

Penelitian ke lapangan adalah suatu cara dalam memperoleh sebuah data yang dilakukan dengan cara melakukan penelitian secara langsung ke lokasi maupun ke objek yang akan diamati. Adapun teknik pengumpulan data yang dilakukan penulis adalah :

##### a. Wawancara (*Interview*)

Wawancara merupakan teknik pengumpulan data dengan cara melakukan tanya jawab secara langsung dengan narasumber yang terkait. Penulis mengadakan wawancara secara langsung dengan salah satu karyawan dari PT.Great Seasons Travel.

#### b. Pengamatan (Observation)

Sebuah metode pengumpulan data yang cukup efektif dalam mempelajari sebuah objek penelitian. Penulis melakukan pengamatan langsung pada PT.Great Seasons Travel.

#### c. Sampel

Merupakan salah satu metode dalam pengumpulan sebuah data untuk mengambil sampel atau contoh-contoh. Penulis meneliti dokumen yang telah tersedia yang memiliki keterkaitan dengan penjualan tiket pesawat di PT.Great Seasons Travel

#### d. Studi Kepustakaan (*Library Research*)

Penulis melakukan studi literatur dalam memperoleh data terkait penulisan skripsi dari berbagai sumber bacaan seperti: buku, jurnal, internet, dan lainnya.

### 2. Prosedur Rancangan

Prosedur ini adalah cara pengumpulan data dengan mempelajari literatur, modul dan manual, buku perpustakaan, dan semua literatur lain yang dianggap perlu dan mendukung.

## 3.3 Operasional Variabel

Variabel adalah segala bentuk apa saja yang dapat ditentukan oleh peneliti untuk dipelajari guna memperoleh informasi dan kesimpulan. Penjelasan variabel dalam penelitian ini adalah sebagai berikut:

#### 1. Jumlah Transaksi

yaitu merupakan nilai yang akan menjadi tolak ukur atau hasil dari sebuah pencapaian. Dalam hal ini transaksi yang terjadi pada setiap bulan sebagai tolak ukur untuk memprediksi peningkatan penjualan. Jumlah transaksi inilah yang

sangat mempengaruhi penelitian ini. Dalam penelitian ini memiliki 6 variabel yaitu jumlah transaksi penjualan pada bulan Mei, Juni, Juli, Agustus, September, Oktober dengan Makapai Lion Air.

### **3.4 Metode Analisis Dan Rancangan Sistem ( *Data Mining* )**

Untuk merancang sebuah sistem Penerapan *Data mining* prediksi penjualan tiket dengan menggunakan metode *autoregressive integrated moving average*, maka perlu dilakukan analisis dan rancangan sistem. Analisis kebutuhan sistem merupakan penjelasan beberapa hal yang menjadi kebutuhan sistem. Berikut adalah kebutuhan sistem dalam perancangan sistem tersebut:

#### 1. Input

Input dalam perancangan sistem penerapan *Data Mining* prediksi penjualan tiket dengan menggunakan metode ARIMA merupakan rekapitulasi transaksi yang dilakukan selama penjualan tiket dalam jangka waktu 6 bulan .

#### 2. Proses

Untuk menentukan tiket maskapai apa dan tujuan mana yang paling banyak dipesan selama proses penjualan terjadi, maka peneliti menggunakan pengimplementasian dengan ARIMA

#### 3. Output

Output atau hasil yang dapat disimpulkan dalam penelitian ini adalah untuk memprediksi peningkatan penjualan tiket maskapai penerbangan.

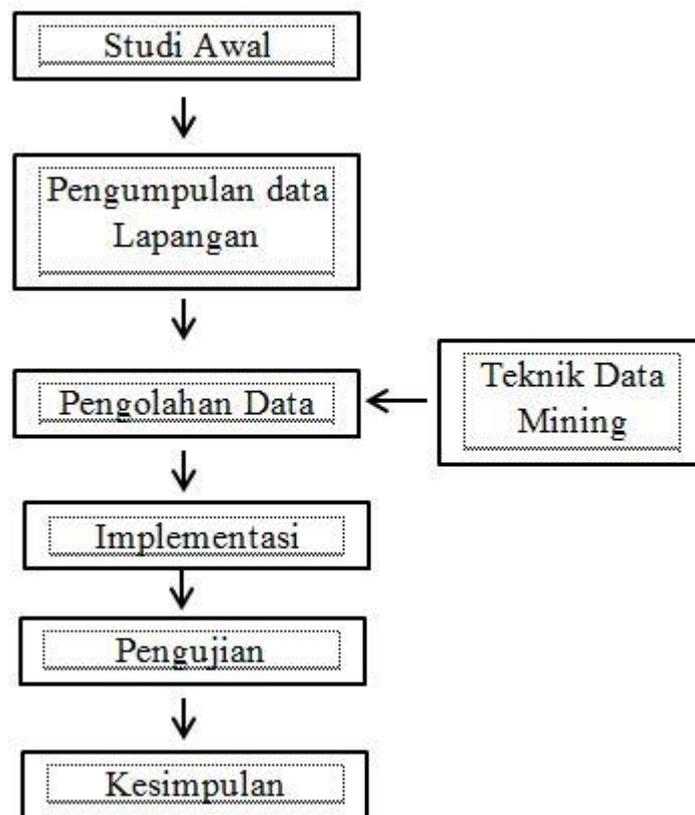
Adapun kebutuhan pendukung dalam proses perancangan sistem tersebut antara lain yaitu :

##### A. *Hardware*

Spesifikasi Kebutuhan perangkat keras dalam penelitian ini adalah sebagai berikut: Laptop

#### B. Software

Spesifikasi Kebutuhan perangkat lunak dalam penelitian ini adalah sebagai berikut: Eviews 9.



**Gambar 3.2** *Flowchart Of Document Analisis Kebutuhan Sistem*

Berikut merupakan tahapan dari *flowchart* penelitian yang telah dibuat:

#### 1. Studi Awal

Langkah pertama dalam penelitian ini adalah menemukan dan menyelidiki masalah yang diteliti. Kemudian menentukan ruang lingkup masalah, latar belakang masalah dari penelitian dan mempelajari beberapa studi literatur yang terkait

dengan masalah dan bagaimana menemukan solusi untuk masalah ini untuk mencapai tujuan yang akan ditentukan. Penting untuk mempelajari bagian dari literatur yang digunakan. Selanjutnya, literatur yang dipelajari sebelumnya dipilih untuk menentukan literatur mana yang digunakan untuk penelitian. Dalam studi literatur ini, teori-teori yang berkaitan dengan penambangan data diperiksa, serta metode autoregresif terintegrasi rata-rata bergerak, yang selanjutnya dibahas.

## 2. Pengumpulan Data

Metode yang digunakan dalam penelitian ini sebagai bagian dari proses pengumpulan data adalah dengan mendokumentasikan file data penjualan yang digunakan sebagai bahan untuk implementasi dan pengujian.

## 3. Pengolahan Data

Pada tahap pengolahan data ini, PT sering dihadapkan pada identifikasi masalah yang ada. Great Seasons Travel dan masalah-masalah yang muncul akan dibahas untuk menemukan solusi. Langkah selanjutnya adalah menganalisis masalah, dengan tujuan agar penulis mengetahui bentuk solusi dan mendapatkan gambaran yang jelas tentang bagaimana mereka dapat diselesaikan. Selain itu, penulis menggunakan teknik data mining dengan metode rata-rata bergerak autoregresif bawaan untuk mendapatkan hasil sebagai tujuan yang ingin dicapai. maka bisa digunakan oleh PT. Great Seasons Travel sebagai pengetahuan untuk meningkatkan penjualan tiket ke publik.

## 4. Implementasi

Pada tahap ini, penulis menggunakan alat atau aplikasi data mining yang ada untuk membantu peneliti memverifikasi hasil analisis yang telah mereka lakukan.

Sementara database digunakan untuk menyimpan data yang diterima, para peneliti menggunakan Microsoft Excel.

## 5. Pengujian

Tahapan ini, pengujian dilakukan menggunakan aplikasi Data Mining dengan menghubungkan sistem yang telah diisi atau yang sudah berisi data yang akan diuji. Mekanisme pengujian sistem dimulai dengan menyiapkan perangkat keras dan perangkat lunak yang akan diperlukan dalam menjalankan sistem yang ada. Sistem yang tersedia adalah aplikasi Data Mining yang pada dasarnya siap dijalankan di komputer.

### **3.5 Lokasi Dan Jadwal Penelitian**

Pada proses penelitian lokasi dan jadwal yang telah ditetapkan sebagai berikut:

#### 1. Lokasi Penelitian

Penelitian dilakukan di PT Great Seasons yang tepatnya berlokasi di Ruko Fanindo Blok O No 15 ( Belakang Kfc ) Tanjung Uncang, Batam.

#### 2. Jadwal Penelitian

Jadwal penelitian ini dilaksanakan sejak Maret sampai Juli 2019, dengan rincian waktu pelaksanaan sebagai berikut :

**Tabel 3.1** Jadwal Penelitian

No	Kegiatan	Maret				April				Mei				Juni				Juli			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Input Judul	■	■	■	■																
2	BAB 1					■	■	■	■												
3	BAB 2									■	■	■	■								
4	BAB 3													■	■	■	■				
5	BAB 4																	■	■	■	■
6	BAB 5																				
7	Penyelesaian																				

Sumber : Data Penelitian ( 2019 )