

**IMPLEMENTASI DATA MINING
PENJUALAN SEPATU MENGGUNAKAN
ALGORITMA APRIORI**

SKRIPSI



Oleh

Wilda Estaria Sitinjak

150210078

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

**IMPLEMENTASI DATA MINING
PENJUALAN SEPATU MENGGUNAKAN
ALGORITMA APRIORI**

SKRIPSI

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



Oleh

Wilda Estaria Sitinjak

150210078

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

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya :

Nama : Wilda Estaria Sitinjak

NPM : 150210078

Fakultas : Teknik

Program Studi : Teknik Informatika

Menyatakan bahwa skripsi yang saya buat dengan judul :

IMPLEMENTASI *DATA MINING* PADA PENJUALAN SEPATU DENGAN MENGUNAKAN ALGORITMA APRIORI

Adalah hasil karya sendiri dan bukan “duplikasi” dari karya orang lain. Sepengetahuan saya, di dalam naskah skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan orang lain, kecuali yang secara tertulis dikutip di dalam naskah ini dan disebutkan dalam sumber kutipan serta daftar pustaka.

Apabila ternyata di dalam naskah skripsi ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah skripsi ini digugurkan dan gelar yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku.

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

Batam, 20 Februari 2020

Wilda Estaria Sitinjak

150210078

**IMPLEMENTASI *DATA MINING*
PADA PENJUALAN SEPATU DENGAN
MENGUNAKAN ALGORITMA APRIORI**

Oleh
Wilda Estaria Sitinjak
150210078

SKRIPSI
Untuk memenuhi salah satu syarat
memperoleh gelar Sarjana

Telah disetujui oleh Pembimbing pada tanggal
seperti tertera di bawah ini

Batam, 20 Februari 2020

Sunarsan Sitohang, S.KOM., M.TI
Pembimbing

ABSTRAK

Penyimpanan data transaksi yang dicatat dan disimpan manual oleh sebuah toko dapat mengakibatkan penumpukan data yang berjumlah besar selama penjualan. Ini akan mengakibatkan data transaksi akan bertumpuk dalam jangka waktu yang cukup lama dan dapat merugikan toko dalam hal kinerja. Di lain sisi, pembeli sering kecewa karena barang yang dicari sudah habis, sementara barang yang lainnya masih memiliki banyak stok di toko dan memiliki jumlah yang tidak seimbang. Penelitian ini memiliki tujuan untuk mempermudah toko Sevendec Shoes dalam menentukan barang yang memiliki tingkat penjualan tinggi, sedang, dan rendah, juga bertujuan untuk melihat dan memantau banyaknya persediaan masing-masing barang yang tersedia guna mengantisipasi pembeli mengalami kekecewaan yang sama. Salah satu upaya yang digunakan adalah sistem penjualan berbasis web yang dirancang guna membantu efektifnya proses penjualan sepatu. Metode yang dipakai pada penelitian ini yaitu algoritma *a priori* yaitu algoritma yang mencari nilai asosiasi barang dan akan menentukan barang apa saja yang memiliki hubungan paling tinggi, sedang, dan rendah. Adapun variabel dalam penelitian ini berupa 14 jenis-jenis sepatu dan data transaksi selama setahun. Pengujian data menggunakan aplikasi Tanagra 1.4. Hasil Akhir dari penelitian ini berupa suatu sistem berbasis web yang dapat me-*record* aktivitas penjualan yang ada, mulai dari data pembelian hingga persediaan barang dan mendapatkan pola aturan asosiasi sebanyak 20 aturan *final* dengan ketetapan nilai minimum *support* sebesar 50% dan nilai *confidence* sebesar 70%.

Kata Kunci : *Data Mining, Association Rule, Tanagra*

ABSTRACT

Storing transaction data that is recorded and stored manually by a store can result in the accumulation of large amounts of data during the sale. This will result in transaction data will accumulate over a sufficiently long period of time and can be detrimental to the store in terms of performance. On the other hand, buyers are often disappointed because the items sought are out of stock, while other items still have a lot of stock in stores and have an unbalanced amount. This study aims to facilitate the Sevendec Shoes store in determining goods that have high, medium, and low sales levels, also aims to see and monitor the amount of inventory of each available item in order to anticipate buyers experiencing the same disappointment. One of the efforts used is a web-based sales system designed to help the effective shoe sales process. The method used in this research is a priori algorithm, which is an algorithm that searches for the value of goods associations and will determine which items have the highest, medium, and lowest relationships. The variables in this study are 14 types of shoes and transaction data for a year. Testing data using the Tanagra 1.4 application. The final results of this study are in the form of a web-based system that can record existing sales activities, ranging from purchasing data to stocking and get an association rule pattern of 20 final rules with a minimum support value of 50% and a confidence value of 70%.

Keywords : *Data Mining, Association Rule, Tanagra*

KATA PENGANTAR

Puji Syukur kehadiran Tuhan Yang Maha Esa yang telah melimpahkan segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata 1 (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 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 selaku Ketua Program Studi Teknik Informatika;
3. Sunarsan Sitohang S.Kom., M.TI selaku pembimbing Skripsi pada Program Studi Universitas Putera Batam;
4. Dosen dan Staff Universitas Putera Batam;
5. Kedua orangtua dan keluarga yang memberikan semangat dan dorongan kepada penulis;
6. Kepada Amelia Sitorus yang membantu dan mendukung secara materi dan non materi;
7. Kepada teman-teman seperjuangan penulis yang telah memberikan masukan dan semangat dalam penyusunan skripsi.

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

Batam, 20 Februari 2020

Wilda Estaria Sitinjak

DAFTAR ISI

SURAT PERNYATAAN ORISINALITAS	ii
ABSTRAK	iv
<i>ABSTRACT</i>	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah	4
1.3 Batasan Masalah	4
1.4 Rumusan Masalah	5
1.5 Tujuan Penelitian	5
1.6 Manfaat Penelitian	6
1.6.1 Aspek Teoritis	6
1.6.2 Aspek Praktis	6
BAB II KAJIAN PUSTAKA	8
2.1 Knowledge Discovery in database (KDD).....	8
2.2 <i>Data Mining</i>	9
2.2.1 Manfaat <i>Data Mining</i>	10
2.2.2 Fungsi <i>Data Mining</i>	11
2.2.3 Teknik Pembelajaran <i>Data Mining</i>	12
2.2.4 Kategori Data Mining.....	12
2.3 Metode <i>Data Mining</i>	13
2.3.1 Klasifikasi Bayes	13
2.3.2 Metode Decision Tree.....	14
2.3.3 Algoritma Nearest Neighbour	14
2.3.4 Fuzzy K-Means	15
2.3.5 Algoritma Apriori.....	16
2.4 Software Pendukung	17
2.4.1 UML	17
2.4.2 PHP.....	22

2.4.3 MySQL	23
2.4.4 StarUML	23
2.4.5 XAMPP.....	24
2.5 Penelitian Terdahulu	24
2.6 Kerangka Pemikiran.....	27
BAB III METODE PENELITIAN	28
3.1 Desain Penelitian.....	28
3.2 Teknik Pengumpulan Data	30
3.2.1 Observasi.....	31
3.2.2 Wawancara.....	31
3.2.3 Dokumentasi	31
3.3 Operasional Variabel.....	31
3.3.1 Data Transaksi (Penjualan).....	31
3.3.2 Jenis Sepatu.....	32
3.4 Metode Analisis dan Perancangan Sistem.....	33
3.4.1 Metode Analisis Perancangan Data Mining	33
3.4.2 Perancangan <i>Data Mining</i>	35
3.4.3 Perancangan Sistem dan Database	35
3.5 Lokasi dan Jadwal Penelitian.....	43
3.5.1 Lokasi Penelitian	43
3.5.2 Jadwal Penelitian.....	43
BAB IV HASIL DAN PEMBAHASAN	45
4.1 Analisis Data.....	45
4.1.1 Seleksi Data.....	45
4.1.2 Menentukan <i>Frequent Itemset</i>	45
4.1.3 Pengolahan Data Satu <i>Itemset</i>	47
4.1.4 Pengolahan Data Dua <i>Itemset</i>	50
4.1.5 Pengolahan Data Tiga <i>Itemset</i>	54
4.1.6 Pembentukan Pola Asosiasi	58
4.2 Implementasi.....	61
4.2.1 <i>Import Data</i>	61
4.2.2 Pemrosesan Data	65
4.3. Halaman Antarmuka Dari Sistem	70
4.4 Pengujian Unit	73

BAB V SIMPULAN DAN SARAN.....	76
5.1 Simpulan	76
5.2 Saran	76
DAFTAR PUSTAKA	77
DAFTAR RIWAYAT HIDUP	
SURAT KETERANGAN PENELITIAN	
SURAT BALASAN PENELITIAN	
LAMPIRAN	

DAFTAR TABEL

Tabel 2.1 <i>Diagram Use Case</i>	18
Tabel 2.2 <i>Diagram Activity</i>	19
Tabel 2.3 <i>Diagram Class</i>	20
Tabel 2.4 <i>Diagram Sequence</i>	21
Tabel 3.1 <i>Jadwal Penelitian</i>	44
Tabel 4.1 <i>Tabel Tabulasi</i>	46
Tabel 4.2 <i>Representasi Data</i>	47
Tabel 4.3 <i>Hasil Perhitungan Itemset 1</i>	49
Tabel 4.4 <i>Aturan Hubungan Itemset 1</i>	49
Tabel 4.5 <i>Hasil Perhitungan Itemset 2</i>	53
Tabel 4.6 <i>Uji Login</i>	73
Tabel 4.7 <i>Tabel Pengujian Input Barang</i>	74
Tabel 4.8 <i>Pengujian Transaksi Barang</i>	74
Tabel 4.9 <i>Pengujian History Transaksi</i>	75
Tabel 4.10 <i>Pengujian Logout</i>	75

DAFTAR GAMBAR

Gambar 2. 1 Tahapan Pada Knowledge Discovery in Database (KDD).....	9
Gambar 2.2 Diagram system berbasis aturan logika.....	16
Gambar 2.3 Kerangka Pemikiran.....	27
Gambar 3.1 Desain Penelitian.....	28
Gambar 3. 2 Rancangan Data Mining.....	33
Gambar 3.3 Analisis Data Mining.....	33
Gambar 3.4 Rancangan Sistem	35
Gambar 3.5 Diagram <i>Use Case</i>	36
Gambar 3.6 <i>Login Diagram Activity</i>	37
Gambar 3.7 <i>Menu Diagram Activity</i>	37
Gambar 3.8 Halaman Utama <i>Diagram Activity</i>	38
Gambar 3.9 <i>Input Barang Diagram Activity</i>	39
Gambar 3.10 <i>Transaksi Barang Diagram Activity</i>	40
Gambar 3.11 <i>History Barang Diagram Activity</i>	40
Gambar 3.12 <i>Logout Diagram Activity</i>	41
Gambar 3.13 <i>Sequence Diagram Proses</i>	42
Gambar 3.14 <i>Class Diagram</i>	43
Gambar 4.1 Tampilan Awal.....	62
Gambar 4.2 Pemilihan <i>Dataset</i>	63
Gambar 4.3 Pemilihan Data.....	63
Gambar 4.4 Data Telah di <i>Import</i>	64
Gambar 4.5 <i>Data View</i>	64
Gambar 4.6 Status Atribut	65
Gambar 4.7 Proses Pemindahan Atribut.....	66
Gambar 4.8 Algoritma <i>Apriori</i>	66
Gambar 4.9 <i>Data Apriori</i>	67
Gambar 4.10 <i>Frequent Itemset</i>	68
Gambar 4.11 Hasil <i>Itemset 2</i>	68
Gambar 4.12 Hasil <i>Itemset 3</i>	69
Gambar 4.13 Seluruh Data <i>Final Rule</i>	69
Gambar 4.14 Tampilan <i>Login</i>	70
Gambar 4.15 Tampilan Awal.....	71
Gambar 4.16 Tampilan Menu <i>Input Barang</i>	71
Gambar 4.17 <i>Transaksi Barang</i>	72
Gambar 4.18 Menu <i>History Transaksi</i>	72

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dunia *fashion* kini semakin maju bahkan mengalami perkembangan yang sangat pesat. Amerika Serikat menjadi acuan dari segi konsumen bagi seluruh negara di dunia ini dalam mengikuti *trend fashion* tersebut (Kompas, 2018). Salah satunya merupakan *trend* sepatu. di Indonesia, sepatu menjadi salah satu *trend* di dunia *fashion* yang dipakai dan disukai banyak kalangan, mulai dari anak-anak sampai dewasa. Setiap orang juga memiliki kriteria sepatu yang mereka suka, ada yang menyukai berdasarkan jenis, model, atau bahkan merek. Selain untuk menunjang penampilan, sepatu juga memiliki manfaat lain, yang mana tergantung kebutuhan masing-masing pemakai. Selain kota metropolitan seperti Jakarta dan Medan, Batam juga menjadi salah satu konsumen sepatu.

Batam adalah satu dari banyak industri terbesar yang berada di Indonesia saat ini. kota sekaligus pulau ini memiliki begitu banyak perusahaan dan tenaga kerja, baik tenaga kerja lokal maupun asing. Seluruh tenaga kerja tersebut tidak pernah luput dari penggunaan sepatu yang digunakan sehari-hari saat bekerja, jalan-jalan, beribadah, atau bahkan berolahraga. Namun, ada sebagian orang gemar mengoleksi sepatu untuk sekedar memenuhi kepuasan tersendiri. Ada banyak tempat penjualan sepatu di kota Batam, salah satunya Toko Sevendec *Shoes* yang menyediakan banyak sekali model dan jenis sepatu yang dicari konsumen. Toko ini juga baru-

baru ini menjual beberapa model tas dengan jumlah yang sedikit. Namun, toko ini lebih membuat titik fokus pada penjualan sepatu.

Persaingan ketat tentunya menjadi tantangan terbesar dalam dunia bisnis apapun, sehingga para pebisnis harus memikirkan taktik dan terobosan baru apa yang akan mereka lakukan kedepannya demi keberlangsungan bisnisnya untuk jangka yang panjang. Salah satunya adalah memanfaatkan data, yaitu data penjualan sepatu. Data penjualan sepatu yang terjual setiap harinya semakin hari semakin banyak dan bertumpuk-tumpuk. Apabila data tersebut tidak diolah dan dibiarkan begitu saja, maka data-data tersebut hanyalah sebatas arsip bagi toko/perusahaan dan kemungkinan akan menjadi sampah. Oleh karena itu, data penjualan harus dimanfaatkan dan diolah menjadi informasi yang penting, serta berguna untuk meningkatkan penjualan sepatu.

Ada beberapa permasalahan bagi toko-toko sepatu yang menjadi tantangan besar untuk tetap bertahan, salah satunya adalah toko sepatu *online*. Dengan adanya toko ini, banyak para konsumen sepatu lebih memilih membeli sepatu secara *online*. Selain karna harganya yang lebih murah, juga lebih simpel karna tidak perlu datang ke toko langsung. Permasalahan kedua, pada saat konsumen ingin membeli sepatu dan menanyakan jenis sepatu yang dicari, penjaga toko akan membutuhkan waktu yang tidak cepat untuk mencarinya dan hal tersebut bisa saja membuat konsumen menunggu lama bahkan membuat kecewa dikarenakan jenis sepatu yang dicari tidak ada/habis. Permasalahan lainnya toko tidak mempunyai informasi yang akurat mengenai tingkat penjualan sepatu per bulan atau per tahun. Data/informasi penjualan sepatu tersebut pada umumnya masih bersifat manual, sehingga para

pemilik toko mengalami kesulitan untuk mengetahui jenis sepatu apa yang paling laku, sedang, dan tidak laku sama sekali, serta jenis sepatu apa yang akan ditambahkan persediaan barangnya di hari-hari berikutnya dan seberapa banyak. Dikarenakan begitu banyak data, sehingga menggunakan algoritma *data mining* yang berguna untuk menganalisa data yang menambah tingkat keakuratan yang tinggi, serta peneliti membuat suatu sistem yang dapat mempermudah penjualan sepatu dengan tak perlu lagi membuat *record* data dengan manual dan dapat melakukan pengecekan banyaknya persediaan barang dengan cepat dan mudah tanpa mengecewakan konsumen yang datang.

Dalam menganalisis data diperlukan sebuah teknik untuk mengetahui perbandingan jumlah jenis sepatu yang terjual sesuai periode yang ditentukan dan melakukan perkiraan. *Data mining* merupakan proses pencarian informasi yang menarik yang ada di dalam data yang terpilih dengan menggunakan algoritma tertentu (Gunadi & Senses, 2012). Data-data yang menjadi hasil dari aktivitas penjualan yaitu data transaksi penjualan sepatu yang diolah menggunakan algoritma *apriori*. Dengan penggunaan algoritma ini, dapat dilihat besarnya dan banyaknya yang terjual berdasarkan prediksi atau hasil yang akan dilakukan serta asosiasi masing-masing jenis sepatu. Penerapan metode *apriori* pada penjualan sepatu dengan jumlah paling banyak terjual. Algoritma *apriori* dipergunakan untuk penggalian informasi, kebenaran lapangan dan pengetahuan serta penambahan dari pengetahuan tersebut (Narti Eka Putria, 2018).

Dengan permasalahan data yang bertambah dari hari ke hari, dapat mengakibatkan data bertumpuk-tumpuk sehingga cara mengolah data tersebut

dibutuhkan penanggulangan yang berkesinambungan (Fauzy, Saleh & Asror, 2016). Data itu lalu membuat suatu bentuk model relasi yang dipakai pada saat membuat perkiraan dan dipakai untuk mencari tau banyaknya sepatu yang laku berdasarkan jenis sepatu. Dari uraian di atas, peneliti tertarik untuk melakukan penelitian yang berjudul **“IMPLEMENTASI DATA MINING PADA PENJUALAN SEPATU DENGAN MENGGUNAKAN ALGORITMA APRIORI”**.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang sudah dipaparkan diatas, maka identifikasi masalah dari penelitian ini adalah :

1. Toko sepatu *online* merupakan salah satu tantangan dalam bisnis penjualan sepatu.
2. Persediaan jenis sepatu yang dicari terkadang habis dan dapat membuat kecewa para konsumen.
3. Sulitnya mengetahui jenis sepatu yang paling dicari dan diminati konsumen, karena data masih bersifat manual.
4. Toko *Sevendec Shoes* masih melakukan *record* data transaksi penjualan secara manual.

1.3 Batasan Masalah

Berdasarkan latar belakang dan identifikasi yang telah dipaparkan di atas, didalam sebuah karya ilmiah diperlukan ada batasan masalah agar menyajikan

permasalahan ini dapat lebih fokus dan terarah. Oleh keterbatasan yang dimiliki peneliti didalam hal kemampuan, dana, waktu dan tenaga maka karya ilmiah ini hanya dibatasi dengan masalah seperti :

1. Peneliti fokus meneliti penjualan sepatu berdasarkan jenis sepatu.
2. Penggunaan data dari periode Januari 2018 selama setahun yang diperoleh dari pemilik toko.
3. Toko sepatu tempat penelitian adalah toko Sevendec Shoes
4. Aplikasi yang digunakan adalah Aplikasi *Tanagra*.
5. Perancangan serta pembuatan sistem menggunakan Bahasa pemrograman *PHP, MySQL*, serta *HTML* berbasis *web*.

1.4 Rumusan Masalah

Berdasarkan batasan masalah yang telah penulis telah pilah, maka rumusan masalah penelitian ini berikut :

1. Bagaimana penerapan algoritma *apriori* untuk mencari tau tingkat penjualan jenis sepatu yang paling banyak terjual?
2. Bagaimana pengujian unit dari sistem berbasis *web*?

1.5 Tujuan Penelitian

Tujuan dari karya ilmiah ini menurut rumusan masalah yang ada, yaitu :

1. Untuk mengetahui hasil dari penerapan algoritma *apriori* dalam penjualan sepatu berdasarkan asosiasi jenis sepatu yang paling banyak terjual.
2. Untuk mengetahui hasil pengujian unit sistem berbasis *web*.

1.6 Manfaat Penelitian

Manfaat penelitian dibagi dua, yaitu :

1.6.1 Aspek Teoritis

1. Menambahkan luas pengetahuan dan ide khususnya peneliti yang membahas *data mining* dengan metode *apriori*
2. Untuk bahan referensi tambahan yang membantu mahasiswa dalam membuat penelitian di masa mendatang.
3. Untuk bahan referensi untuk pembaca di dalam membuat suatu penelitian.

1.6.2 Aspek Praktis

1. Bagi toko *Sevendec Shoes*
Membantu pemilik toko *Sevendec Shoes* dalam mengolah data yang berjumlah cukup besar serta dengan adanya sistem yang telah dirancang oleh peneliti, pemilik/penjaga toko dapat lebih mudah melakukan pemeriksaan persediaan sepatu dengan cepat dan mudah, serta tidak perlu melakukan *record* manual ke dalam buku catatan biasa.
2. Bagi Akademik.
Sebagai sumber tambahan dalam membantu mahasiswa melakukan penelitian selanjutnya dan sumbangan pemikiran dalam melakukan prediksi penjualan jenis sepatu dengan menggunakan *data mining* algoritma *apriori*.
3. Bagi Peneliti

Dapat menambahkan wawasan ilmu dan membuat keahlian semakin meningkat untuk menerapkan *data mining* algoritma *apriori* dalam memprediksi penjualan

sepatu berdasarkan jenis di toko *Sevendec Shoes*. Selain itu, peneliti juga mendapatkan kesan yang baru didalam menyelesaikan karya ilmiah ini yang dapat berguna bagi tempat penelitian

BAB II

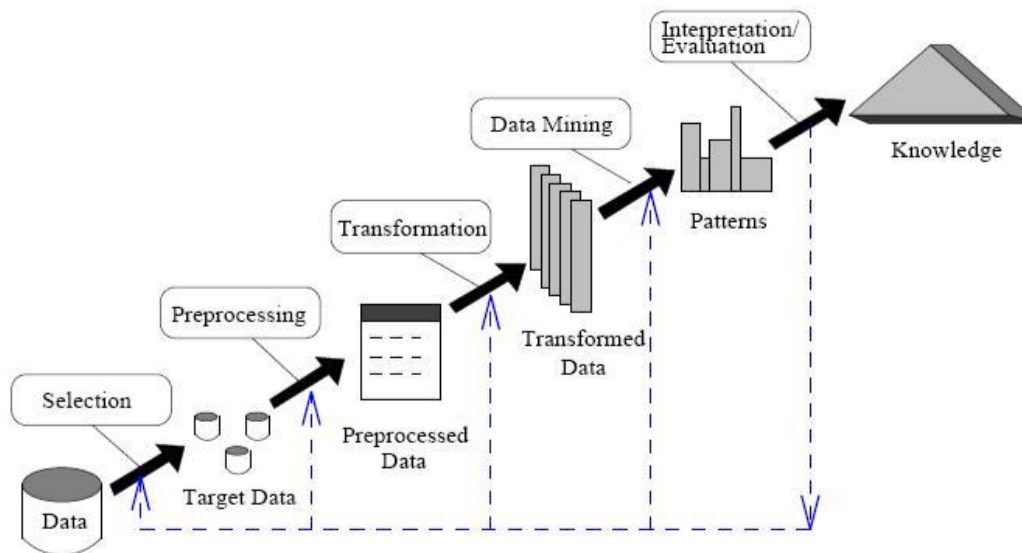
KAJIAN PUSTAKA

2.1 *Knowledge Discovery in database (KDD)*

Knowledge Discovery in Database merupakan pencarian pengetahuan yang ada pada database yang sangatlah besar. Salah satu rangkaian *Knowledge Discovery in Database* merupakan *Data Mining*. KDD memiliki ikatan dengan sebuah algoritma temuan serta integrasi, visual, dan interpretasi dari sejumlah data. Rangkuman algoritma ini memiliki tahap yaitu (Retno Tri Vlandari, 2017, hal:2-3):

1. Membersihkan data (pembuangan data).
2. Integrasi data (gabungan data).
3. Bersifat transformal, yaitu dapat mengubah data mejadi data diolah ke dalam *data mining*.
4. Teknik *data mining*, proses ekstraksi data.
5. Evaluasi pola (proses menjadi pengetahuan dan bisa dipergunakan sebagai pendukung diambilnya suatu keputusan).
6. Presentasi pengetahuan.

Seluruh tahap di atas adalah bagian dari algoritma karya ilmiah pengetahuan yang terdiri dari tinjauan suatu informasi maupun pola itu tidak berlawanan dengan hipotesa yang ada. Lalu dari KDD yaitu wawasan dijabarkan kedalam bentuk dan mudah dimengerti oleh *user*.



Gambar 2. 1 Tahapan Pada Knowledge Discovery in Database (KDD)
Sumber : Retno Tri Vulandari, 2017

Istilah *data mining* dan *Knowledge Discovery in Database* seringkali digunakan secara bergantian untuk memahami proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain.

2.2 Data Mining

Data mining adalah rangkaian tahapan penggalian nilai kelebihan berbentuk info yang tidak diketahui baik secara manual dari sistem *database*. Informasi didapat dengan melakukan ekstraksi dan mengetahui bentuknya yang menarik dari data yang ada pada *database* (Retno Tri Vulandari, 2017).

Perspective dan wawasan yang tidak sama membuat para ahli dan pakar memiliki pengertian yang tidak sama pada *data mining*. *Data Mining* merupakan kegiatan menambang data-data untuk penggalian informasi serta pengetahuan yang berharga. Pengetahuan ini dapat berupa pola data ataupun relasi antar data-data

yang *valid* atau data yang tidak diketahui. *Data Mining* atau penambangan data adalah teknik yang instan untuk menemukan pengetahuan, pola dan atau relasi antar data, secara otomatis (Suyanto, 2017).

2.2.1 Manfaat *Data Mining*

Dua sudut pandang untuk memanfaatkan *data mining* yaitu dari *perspective* komersil dan keilmuan (Retno Tri Vulandari, 2017).

1. Dari *perspective* komersil, *data mining* dapat dipergunakan dalam penanganan meledaknya isi data, dengan komputasi, yaitu guna mendapatkan informasi yang bersifat penting dan ternyata aset dalam memicu daya saing meningkat.

Contohnya :

- a. Bagaimana mencari tau konsumen yang hilang karena pesaing banyak.
- b. Bagaimana mencari tau *product* dan atau konsumen yang mempunyai karakter yang sama.
- c. Bagaimana mengenali *product* yang dibeli secara bersama dengan produk lain.
- d. Bagaimana membuat suatu prediksi dari penjualan.
- e. Bagaimana mengetahui akibat didalam penentuan banyaknya produksi *item*.
- f. Bagaimana memperkirakan *behaviour* bisnis di masa mendatang.
- g. Dari segi keilmuan, *data mining* dipergunakan menangkap, membuat analisis dan melakukan simpan data yang *realtime* dan besar.
- h. *Remote sensor* terdapat di luar angkasa.
- i. *Telescope* yang dipergunakan di angkasa.

- j. Simulasi saintifik yang dapat meningkatkan data dengan ukuran yang besar atau setara dengan *size terrabytes*.

2.2.2 Fungsi Data Mining

Kegunaan atau fungsi *data mining* yang biasa dan sering diimplementasikan yaitu (Retno Tri Vulandari, 2017, hal:4-5):

1. *Association*, yaitu tahap mencari dan mendapatkan *rule* asosiasi antara perpaduan *item*.
2. *Sequence*, proses untuk mencari *rule* hubungan antar kombinasi *item*, dan diimplementasikan lebih dari satu periode.
3. *Clustering*, merupakan proses pengelompokkan data yang bisa dibilang sama.
4. *Classification*, proses temuan dari tipe yang dapat menjelaskan atau membedakan konsep dari suatu kelas data, dengan tujuan membuat suatu perkiraan dari kelas obyek belum diketahui.
5. *Regression*, merupakan suatu pemetaan data yang ada pada suatu perkiraan.
6. *Forecasting*, yaitu tahap melakukan estimasi perkiraan bernilai berlandaskan polanya yang ada pada kumpulan data.
7. *Solution*, merupakan tahap terakhir temuan dari *basic* suatu permasalahan dan pemecahan masalah dan dihadapkan atau sebagai informasi dalam proses pengambilan putusan akhir.

2.2.3 Teknik Pembelajaran *Data Mining*

Teknik pada data mining terdiri dari 3 metode pelajaran, yaitu (Retno Tri Vlandari, 2017, hal:7-8)

a. *Supervised Learning*

Supervised Learning adalah metode yang sangat banyak digunakan. Metode ini mempengaruhi tahapan berlatih yang berada di riwayat pelatihan dan memetakan karakter menjadi sesuatu yang dikenali serta melakukan pengolahan ke metode pada *data mining*.

b. *Unsupervised Learning*

Metode ini tidaklah mempengaruhi tahapan berlatih, namun ketergantungan dengan menggunakan algoritma yang dapat mengenali pola, seperti *association*, serta *sequence*, yang timbul dari ciri-ciri spesifikasi pada *input* data.

c. *Reinforcement Learning*

Metode ini tidaklah sering dipergunakan jika dibanding dengan metode-metode di atas, namun mempunyai terapan yang selalu melakukan optimasi yang bersifat *time to time* serta mempunyai penanganan yang bersifat adaptisasi.

2.2.4 Kategori *Data Mining*

Data mining terbagi dua kategori (Retno Tri Vlandari, 2017).

a. Bersifat Prediksi

Kategori sifat ini bertujuan agar dapat melakukan prediksi nilai atribut sesuai nilai atribut lainnya. Biasanya ini dikenal sebagai target atau pun variabel tak bebas.

b. Bersifat Pendeskripsian

Kategori ini bertujuan agar dapat mengurangi dan mempersingkat dari suatu relasi dan bersifat inti didalam data. Pekerjaan ini biasanya melakukan invesitasi serta sering butuh terhadap suatu metode agar hasil terakhir menjadi *valid* dan jelas.

2.3 Metode Data Mining

Pada data minig, terdapat metode yang cukup membantu dalam permasalahan olahan data, yaitu :

2.3.1 Klasifikasi Bayes

Metode klasifikasi Baye yaitu metode dan bersifat pada fundasi statistik di *data mining*. Metodenya berlandaskan kuantifikasi *trade-off* antara aneka pengklasifikasian *rule* pada *probability* (peluang) (Retno Tri Vulandari, 2017). Algoritma ini pertama kali dicetus dan dibuat Thomas Bayes pada abad ke-18. Disini, probabilitas yang memiliki syarat ditulis sebagai

$$P(H|X) = \frac{P(X|H)P(H)}{P(H)}$$

Rumus 2.1 *Probability* Algoritma Bayes

Penjelasan formula berikut yang mana X berperan menjadi bukti, H menjadi hipotesa, $P(H|X)$ menjadi peluang di mana hipotesa H bernilai tepat untuk bukti X atau sinonim $P(H|X)$ yaitu peluang *posterior* H dengan syarat X, $P(H|X)$ merupakan probability X tepat menjadi hipotesa H atau probability *pasterior* X dengan kriteria H, $P(H)$ yaitu probabilitas yang memprioritaskan hipotesa H, dan $P(X)$ merupakan probability yang memprioritaskan X (Suyanto, 2017, hal:126-129).

2.3.2 Metode *Decision Tree*

Metode ini berupa sebuah pengklasifikasian praktikal dan menjadi yang terbanyak dipergunakan juga sangatlah populer (Suyanto, 2017). Pohon keputusan analisa memecahkan sebuah permasalahan merupakan taktik dengan memetakan bagaimana memecahkan suatu permasalahan yang didapat pada diri permasalahan tersebut. Skema ini memperlihatkan akibat probabiliti yang mempengaruhi jalan lain dari pengambilan putusan, serta prediksi hasil akhir yang didapat jika mengambil alternative pada keputusan (Retno Tri Vuldari, 2017).

2.3.3 Algoritma *Nearest Neighbour*

Situasi di mana ada banyak data pencilan yang menyebabkan representasi model berbasis aturan, model distribusi kepadatan peluang ataupun model klasifikasi yang sulit dihasilkan. Metode atau model klasifikasi ini dapat digunakan dan tidak memerlukan parameter yang rumit. Metode ini dikenal dengan *Nearest Neighbour Rule*. Istilah rule disini bukan memperlihatkan metode ini sebagai klasifikasi yang berbasis pada aturan, tetapi mengatakan *rule* untuk menentukan tetangga paling dekat berdasarkan ukuran statistik (Suyanto, 2017). Adapun metode ini terbagi dua yaitu :

1. *k-Nearest Neighbour Rule*

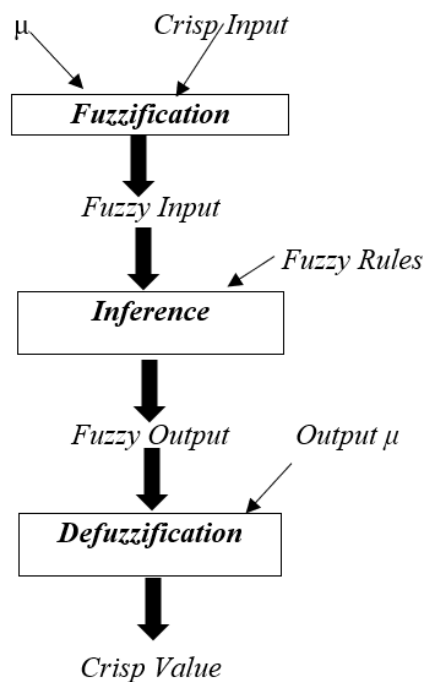
Metode ini mencari lalu menemukan beberapa k objek data ataupun pola terdekat dengan pola masukan, lalu memilih kelas dengan jumlah pola paling banyak diantara pola tersebut. Jika dicontohkan, metode ini melakukan klasifikasi pola dengan *voting*.

2. *Pseudo Nearest Neighbour Rule*

Pseudo Nearest Neighbour Rule (PNNR) merupakan versi terbaru dari kNNR yang bekerja pada local dan internal. PNNR dirancang untuk mengatasi kelemahan metode kNNR yang pada umumnya memberikan cara kerja yang rendah pada data yang derau. PNNR juga dapat digunakan untuk menghitung jarak total pola masukan yang tak berlabel dengan beberapa pola terdekat lalu memutuskan kelas dengan jumlah jarak yang paling minimal sebagai kelas keputusan pada pola tersebut (Suyanto, 2017).

2.3.4 Fuzzy K-Means

K-means adalah metode pengklasifikasian dan dilakukan dengan mengulanginya. Metode ini mengambil nilai *cluster* acak, dan menjadikannya sebagai titik acuan atau *centroid*, *mean*, atau *means*. Lalu menghitung jaraknya, ke setiap *centroid* dengan memakai rumus *Euclidian* sampai akhirnya menemukan jarak terdekat dengan data *centroid*. Metode ini disebut juga dengan *segmentation* yang mana digunakan untuk melakukan identifikasi kelompok alami dari pada kasus yang didasarkan pada suatu atribut, mengelompokkan data yang memiliki kemiripan atribut (Retno Tri Vulandari, 2017).



Gambar 2.2 Diagram system berbasis aturan logika
Sumber : Data Olahan Penelitian, 2019

2.3.5 Algoritma Apriori

Algoritma apriori merupakan tahapan pada proses untuk mendapatkan jumlah *set* data, dan mengiterasikan dengan data (Fauzi, Saleh & Asror, 2016). Analisa asosiatif merupakan metode pada *data mining* agar mendapatkan *rule* hubungan atau asosiasi antar gabungan *item*. Contohnya menganalisa saat membeli di suatu warung dengan mengetahui seberapa tinggi tingkat mungkin pembeli A membelikan mie serentak dengan membeli telur (RetnoTri Vlandari, 2017). Ada dua dasar analisis asosiasi yaitu :

1. Analisis pola dengan jumlah yang tinggi

Proses untuk menemukan gabungan *item* yang melampaui persyaratan minimal dari *support* data. Nilai dicari menggunakan formula :

$$\text{support } A = \frac{\text{Jumlah transaksi mengandung } A}{\text{Total transaksi}}$$

Rumus 2.2 Mencari nilai Support A

nilai *support* dari dua *item* didapatkan dari formula :

$$\text{support } A \cap B = \frac{\text{Jumlah transaksi mengandung } A \text{ dan } B}{\text{Total transaksi}}$$

Rumus 2.3 Mencari gabungan Support A irisan B

- a. Membentuk *rule* asosiatif

Seluruh pola pada frekuensi tinggi telah didapatkan, cari *rule* asosiatif dan melampaui syarat minimal untuk *confidence* dan mencari *confidence* dengan rumus :

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

Rumus 2. 4 Rumus *confidence*

2.4 Software Pendukung


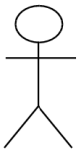



2.4.1 UML

Pemodelan yang lebih sering dipakai saat ini yaitu UML (*Unified Modelling Languages*). UML merupakan satu dari begituu banyak bahasa yang standar dan dipergunakan pada perindustrian supaya mengilustrasikan *requirement*, menciptakan analisa & rancangan, mendefinisikan dan mendeskripsikan arsiteknya. UML juga merupakan suatu standarisasi bahasa pemodelan yang muncul dikarenakan ada pemodelan visualisasi yang dibutuhkan agar melakukan spesifikasi, penggambaran, pembangunan, dan pendokumentasian *software* (Rosa, M. Shalahuddin, 2013).

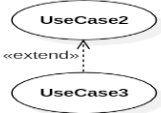
1. Use Case Diagram

Use Case Diagram adalah *modelling* untuk kelakuan dari sistem informasi yang hendak dibentuk. *Use Case* menggambarkan suatu aksi antar yang satu atau lebih dengan sistem informasi yang hendak dibentuk. *Use Case* biasanya dipergunakan agar mencari tau apa kegunaan pada suatu sistemasi serta orang yang memiliki hak untuk dapat mempergunakannya.

Tabel 2.1 Diagram Use Case

Simbol	Deskripsi
<p><i>Use Case</i></p> 	<p>Fungsi yang tersedia oleh sistem untuk beberapa unit dan bertukaran pesan antara unit/aktor; sering dan dapat digambarkan melalui pemakaian kata kerja di frase awal <i>use case</i>.</p>
<p>Aktor</p> 	<p>Dalam hal ini Aktor lebih sering dinyatakan sebagai orang yang melakukan interaksi dengan sistem informasi. Biasanya digambarkan memakai kata benda pada frase awal.</p>
<p>Asosiasi</p> 	<p>Komunikasi antar <i>actor</i> dan <i>use case</i> dan memiliki partisipasi pada <i>use case</i> akan mempunyai interaksi dan hubungan dengan aktor.</p>
<p>Ekstensi/<i>Extend</i></p> 	<p><i>Use case</i> tambahan dapat berdiri sendiri walau tanpa <i>use case</i> tambahan tersebut. Mirip dengan prinsip <i>Inheritance</i> pada pemrograman berorientasi objek. Yang menjadi tambahan yaitu nama depan sama dengan <i>use case</i> yang ditambahkan.</p>
<p>Generalisasi</p> 	<p>Asosiasi umum-khusus, memiliki fungsi sesuai dengan nama masing-masing.</p>

Tabel 2.1 Diagram *Use Case* (Lanjutan)






<i>Include</i>	Hubungan <i>Use Case</i> tambahan ke dalam sebuah <i>Use Case</i> memerlukan dirinya sendiri agar berfungsi.
<p><i>Extend</i></p> 	Pada ekstensi ini, diambil hanya induk <i>Use Case</i> dan dijadikan kelas dengan metode ekstensinya.

Sumber : Rosa, M. Shalahuddin, 2013


2. Activity Diagram

Activity Diagram ataupun diagram aktivitas mendeskripsikan proses dari sistem berbasis dan menu yang tersedia di perangkat lunak. Pendeskripsian diagram aktivitas sistem bukanlah berdasarkan aktivitas *actor*, melainkan aktivitas dari sistem itu.

Tabel 2.2 Diagram *Activity*

Simbol	Deskripsi
Status awal 	Status awal aktivitas, semua sistem selalu diawali dengan simbol ini.
Aktivitas 	Aktivitas yang dilakukan sistem, diawali dengan kata kerja.
Percabangan 	Asosiasi percabangan di mana apabila ada pilihan aktivitas yang lebih dari satu.
Penggabungan/join 	Asosiasi penggabungan di mana jika lebih dari satu aktivitas maka digabung.
Status akhir 	Status akhir dilakukan oleh sistem, selalu diakhiri dengan status akhir.

Tabel 2.2 Diagram *Activity* (Lanjutan)

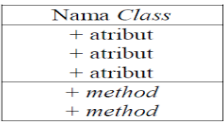



Swimlane 	Pemisahan organisasi-organisasi, serta menanggungjawab terhadap kegiatan saat terkini.
--	--

Sumber : Rosa, M. Shalahuddin, 2013




3. *Class* Diagram

Class Diagram atau diagram kelas mengilustrasikan bentuk suatu sistem dari sudut definisian kelas dimana dibuat untuk membangun sistem. Kelas mempunyai atribut dan metode, dimana atribut tersebut adalah variabel-variabel yang ada pada sebuah kelas tertentu dan merupakan kegunaan yang dimiliki oleh suatu kelas.

Tabel 2.3 Diagram *Class*

Simbol	Deskripsi
	Kelas pada bentuk sistem
Antarmuka  nama_antarmuka	Sama dengan konsep <i>interface</i> dalam pemrograman berorientasi objek.
Asosiasi 	Hubungan antarkelas dengan artian pada umumnya.
Asosiasi berarah 	Hubungan antar kelas dengan artian kelas yang satu dipergunakan kelas lainnya.

Tabel 2.3 Diagram *Class* (Lanjutan)



General 	Hubungan antar kelas dengan arti general-spesial (umum-khusus).
<i>Independency</i> 	Hubungan antar kelas dengan arti tergantung antar kelas.
Agregatif 	Hubungan antar kelas pada arti semua-bagian (<i>whole part</i>)

Sumber : Rosa, M. Shalahuddin, 2013

4. *Sequence* Diagram






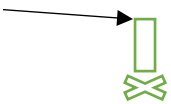
Diagram sekuen mendeskripsikan tingkah laku suatu *object* dalam *use case* dengan waktu hidup objek serta pesan yang terkirim dan masuk antara objek. Skenario pada *use case* dibutuhkan dalam membuat diagram sekuen ini dan seharusnya mengidentifikasi objek terkait di dalamnya.

Tabel 2.4 Diagram *Sequence*

Simbol	Deskripsi
 Nama aktor	Pelaku, tahapan, atau sistem lainnya yang berinteraksi dengan <i>system</i> informasi yang hendak ada di luaran <i>system</i> .
Garis hidup / <i>lifeline</i> 	Mendefinisikan hidup sebuah objek.
Objek <div style="border: 1px solid black; padding: 2px; width: fit-content;">nama_objek:nama_kelas</div>	Mendefinisikan objek yang berinteraksi pesan.

Sumber : Rosa, M. Shalahuddin, 2013

Tabel 2.4 Diagram *Sequence* (Lanjutan)

<p>Waktu aktif</p> 	<p>Objek dalam keadaan aktif dan berinteraksi, yang terhubung dengan waktu aktif ini adalah tahapan yang dilakukan.</p>
<p>Pesan tipe <i>create</i> <<create>></p> 	<p>Menyatakan suatu objek membuat objek yang lain, arah panah mengarah pada objek yang dibuat.</p>
<p>Pesan tipe <i>call</i> 1 : nama_metode()</p> 	<p>Menyatakan suatu objek untuk memanggil operasi/metode yang ada pada objek lain atau diri sendiri.</p>
<p>Pesan tipe <i>send</i> 1: masukan</p> 	<p>Menyatakan bahwa suatu objek mengirimkan data/masukan/informasi ke objek lainnya, arah panah mengarah pada objek yang dikirim.</p>
<p>Pesan tipe <i>return</i> 1: keluaran</p> 	<p>Menyatakan bahwa objek yang telah menjalankan suatu operasi menghasilkan suatu kembalian ke objek tertentu, arah panah mengarah pada objek yang menerima kembalian.</p>
<p>Pesan tipe <i>destroy</i> <<destroy>></p> 	<p>Menyatakan suatu objek mengakhiri hidup objek yang lain, arah panah mengarah pada objek yang diakhiri, sebaiknya jika ada <i>create</i> maka ada <i>destroy</i>.</p>

Sumber : Rosa, M. Shalahuddin, 2013

2.4.2 PHP

PHP adalah *programming script* pada sisi penyedia suatu web, script membuat dokumennya HTML langsung, dengan maksimal HTML yang dihasilkan dari aplikasi bukan dokumen HTML yang dibuat memakai editan *text* ataupun *editor* HTML PHP/FI merupakan nama awal dari PHP. PHP merupakan rumah

pribadi, sedangkan FI ialah antarmuka bentuk. PHP pertama kali diciptakan oleh Rasmus Lerdoff. PHP menjadi program yang bersifat khusus untuk menerima masukan melalui formulir yang diterima dari dalam *web browser*. *Script* PHP dikelilingi dengan menggunakan tag awal dan tag akhir khusus (tag awal `<?php` atau `<?>` dan tag akhir `?>`), pemrograman wajib untuk keluar masuk dari mode script PHP. Pekerjaan dasar dari PHP adalah mendapatkan data dari *form*, menghasilkan isi web yang dinamik serta menerima *cookies*, namun fitur yang terpercaya adalah mendukung banyak *database*. (Febio, 2011, pp. 43-44)

2.4.3 MySQL

MySQL merupakan basis data yang cuma mengoperasikan *Crash Unix day Linux*, yang mana SQL merupakan singkatan dari *Structured Query Language*. SQL ialah sebuah *request* bahasa yang teratur dan terorganisir dan menempel pada satu *data base*, dan MySQL merupakan basis data nya. SQL adalah perintah yang menempel didalam SMBD. Adapun struktur dasar dari *express* SQL terdiri dari tiga klausa, yaitu : *select*, *from*, dan *where*. (Febio, 2011,p. 44).

2.4.4 StarUML

UML (Unified Modelling Language) adalah Bahasa yang memiliki notasi lengkap hingga menjadi suatu ketentuan bersifat visual. UML rilis pada tahun 1994, dan tokoh dari penemu UML saat itu Booch, Runbaugh, dan Jacobson. Adapun tujuan UML yaitu memberikan modal berorientasi objek yang siap digunakan dan memberikan bahasa yang dipergunakan manusia dalam pemodelan mesin (Yasin, 2012).

2.4.5 XAMPP

Pada perangkat lunak ini peneliti menggunakan Xampp versi 3.2.2 untuk *running web server Apache, MySQL* dan *FileZilla* sebagai penyedia basis data dan pencarian *web* (Sandy Kosasi, 2014)

2.5 Penelitian Terdahulu

Beberapa penelitian terdahulu digunakan penulis sebagai referensi dalam membuat penelitian, antara lain :

1. Penelitian Dewi Kartika Pane dengan judul “Implementasi data mining pada penjualan produk elektronik dengan algoritma *apriori* (Studi Kasus : kreditplus)” menjelaskan bahwa data penjualan di kreditplus selama ini berantakan dan tidak teratur, hingga data tersebut hanya berguna menjadi arsip untuk perusahaan dan tidak berguna lagi untuk mengembangkan strategi pemasaran. Untuk itu peneliti membuat suatu analisis untuk memprediksi laptop dengan *merk* apa yang paling banyak terjual agar tidak terjadi penumpukan barang/laptop kedepannya. Adapun peneliti menggunakan *tool* Tanagra.
2. Penelitian Mohamad Fauzi, Kemas Rahmat Saleh, Ibnu Asror yang berjudul “Penerapan metode *Assosiation Rule* menggunakan algoritma *apriori* pada simulasi prediksi hujan wilayah kota Bandung” menjelaskan bahwa prediksi cuaca merupakan hal yang dibutuhkan orang di dunia. Saat melakukan perkiraan hujan, mengolah data cuaca merupakan hal penting. Akan tetapi masalahnya, data cuaca tersebut bertambah setiap harinya, dan perkiraan cuaca dibutuhkan untuk keesokan harinya. Oleh karena itu dibutuhkan metode *data mining* yang berakurasi tinggi

untuk menyelesaikannya. Hasil akhir dari penelitian ini berupa aturan-aturan asosiasi, dimana aturan tersebut dapat dijadikan sebagai acuan dalam memprediksi cuaca hujan atau tidak hujan untuk satu hari kedepan.

3. Penelitian Arif Ismail Husin, Farida Mulyaningsih yang berjudul “Penerapan metode data mining analisis terhadap data penjualan pakaian dengan algoritma *apriori*” menjelaskan jumlah penjualan pakaian khususnya remaja mengalami peningkatan setiap bulannya. berbagai macam merk pakaian yang ditawarkan membuat pembeli makin tertarik agar belanja. Untuk mengetahui merk pakaian remaja dengan penjualan terbanyak dengan serentak dibutuhkan suatu Teknik agar mudah dalam mendapat informasi seputar data penjualan. metode yang digunakan ialah metode *apriori* dan dalam mengolah data penelitian digunakan *tool* Tanagra. Dengan algoritma *apriori* perusahaan dapat membentuk taktik dengan mempromosikan produk yang paling laku dengan produk lain yang akan meningkatkan penjualan setiap bulannya.

4. Penelitian Narti Eka Putria yang berjudul “*Data Mining* penjualan tiket pesawat menggunakan algoritma *apriori* pada terminal tiket Batam tour & travel” menerangkan semakin majunya teknologi informasi saat ini semakin diperlukan juga tingkat kreativitas dan inovasi agar penjualan produk meningkat. Hasil dari penelitian ini menerapkan algoritma *apriori* untuk melihat pola pembelian konsumen saat pengolahan data dapat diketahui tiket yang paling sering dipesan pembeli, sehingga pihak travel dapat mengatur strateginya agar tiket dapat terjual lebih banyak lagi.

5. Penelitian Ayuni Asistiyasari, Taufik Baidawi dalam judul “Analisis penerimaan karyawan posisi *Field Collector* menggunakan algoritma C4.5 pada PT. Primas Jamintara Jakarta” menjelaskan bahwa proses penerimaan karyawan, khususnya *Field Collector* melalui tahap seleksi penilaian kriteria-kriteria seperti usia, pendidikan, pengalaman, dan kepemilikan kendaraan. Hal ini dilakukan agar mendapatkan kriteria yang dibutuhkan oleh pelanggannya. Penggunaan algoritma C 4.5 untuk menganalisis data dalam jumlah yang banyak menggunakan konsep *entropy* dan *gain*.

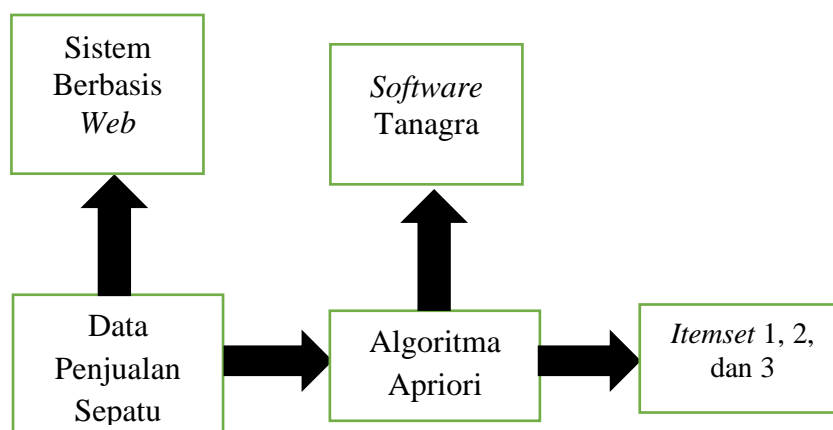
6. Penelitian Jiao Yabing yang berjudul “Research of An Improved Algorithm In Data Mining Association Rules” *explain that the classic algorithm of association rules, which enumerate all of the frequent item sets. When this algorithm encountered dense data due to the large number of long patterns emerge, this algorithm's performance declined dramatically. Finally, the improved algorithm is verified, the results show that the improved algorithm is reasonable and effective, can extract more value information.*

7. Penelitian Abdar da Arji dengan title “Comparing Performance of Data Mining Algorithms in Prediction Heart Diseases” *to predict the risk of heart diseases. After feature analysis, models by five algorithms including C5.0, Neural Network, Support Vector Machine (SVM), K-Nearest Neighborhood (KNN) and Logistic Regression, developed and validated. C5.0 Decision tree has been able to build a model with greatest accuracy 93.02%, KNN, SVM, Neural Network have been 88.37%, 86.05% and 80.23% respectively. Produced results of decision tree*

can be simply interpretable and applicable; their rules can be understood easily by different clinical practitioner.

2.6 Kerangka Pemikiran

Untuk memprediksi penjualan sepatu, dibutuhkan analisa data. Algoritma *apriori* adalah Teknik tersering dipergunakan dalam memprediksi penjualan atau pun menganalisi data pada tingkat akurasi yang sangat tinggi. Data yang akan dianalisa merupakan data transaksi sepatu, yang mana data akan dilakukan pengolahan dengan memakai perangkat Tanagra agar mendapat perkiraan dengan akurasi tinggi. Dari rincian latarbelakang dan algoritma yang digunakan, maka kerangka pemikiran penelitian sebagai berikut :



Gambar 2.3 Kerangka Pemikiran
Sumber : Data Olahan Penelitian, 2019

BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Dalam melakukan sebuah penelitian, pasti dibutuhkan yang namanya perencanaan atau perancangan, agar penelitian tersebut dapat berjalan secara sistem. Sebuah perencanaan atau perancangan penelitian yang membahas tentang *data mining* ini akan direpresentasikan ke dalam sebuah desain penelitian yang akan dirangkum dan dijelaskan dengan melalui gambar di bawah ini :



Gambar 3.1 Desain Penelitian
Sumber : Data Olahan Penelitian, 2019

Tahapan atau proses dari Gambar 3.1 di atas yang akan dijelaskan dalam penelitian ini sebagai berikut :

1. Identifikasi Masalah

Karya Ilmiah ini diawali dengan studi dulu, studi memiliki tujuan untuk memilih permasalahan yang berhubungan pada judul penelitian, sehingga peneliti mengetahui cara agar dapat menyelesaikan permasalahan tersebut.

2. *Studi Literature*

Dalam *fase* ini, peneliti membaca dan belajar dari buku, jurnal, internet dan pernyataan yang memiliki kaitan dengan yang diteliti.

3. Pengumpulan Data

Melakukan *interview* (wawancara) dengan pemilik dari toko *Sevendec Shoes* serta melakukan observasi ke toko langsung, supaya mendapat data penjualan sepatu dan akan dibuat sebagai sumber data untuk penelitian.

4. *Data Mining* dengan algoritma *Apriori*

Setelah data-data telah terkumpul, selanjutnya adalah menganalisa data dengan menggunakan algoritma *apriori* atau *rule association*.

5. Implementasi dengan *Tanagra*

Setelah data-data yang sudah terkumpul dianalisa dan menghasilkan *output* berupa penjualan sepatu berdasarkan jenis, maka selanjutnya dilakukan implementasi dengan *Tanagra* untuk melakukan pengujian dari data penjualan sepatu yang dihitung manual dengan data yang berdasarkan pengujian sistem *Tanagra*.

6. Desain Sistem

Pada tahapan ini, peneliti mendesain suatu sistem penjualan sepatu di toko *Sevendec Shoes*, dikarenakan toko selama ini melakukan perhitungan, pengecekan stok dan catatan manual dalam hal penjualan sepatu.

7. Sistem Berbasis *Web*

Setelah tahap desain sistem selesai, maka hasilnya adalah berupa penjualan sepatu dengan sistem berbasis web.

8. Hasil

Setelah seluruh proses diatas telah dibuat, akan didapat sebuah hasil berupa *itemset* dengan jumlah penjualan terbanyak. Maka terbentuklah suatu *rule assosiasi* dari penjualan sepatu, dan sebuah sistem berbasis *Web* yang dapat menunjang kinerja di toko menjadi lebih mudah dan efisien.

3.2 Teknik Pengumpulan Data

Pengumpulan data adalah sebuah metode yang digunakan untuk mendapatkan data yang dibutuhkan dalam suatu penelitian yang berasal dari narasumber yang bersangkutan. Berdasarkan sumber datanya, mengumpulkan suatu data dilakukan dengan sumber primer dan sekunder. Sumber primer merupakan sumber data yang diberikan langsung oleh narasumber kepada pengumpul data. Sedangkan sumber sekunder merupakan sumber data yang diberikan secara tidak langsung oleh narasumber kepada pengumpul data dapat dilakukan lewat perantara, seperti dokumen (Prof. Dr. Sugiyono. 2014; 224). Berdasarkan cara, maka teknik pengumpulan data dapat dilakukan dengan observasi, wawancara, kuesioner, dokumentasi, dan gabungan.

3.2.1 Observasi

Observasi dilakukan peneliti dengan datang langsung ke toko sepatu *Sevendec Shoes*, dan melihat jenis-jenis sepatu yang ada disana ada sebanyak 14 jenis. *Size* sepatu juga berada pada *range* 37 - 40 untuk dewasa, dan 27 – 32 untuk anak-anak.

3.2.2 Wawancara

Wawancara juga diadakan peneliti dengan pemilik toko dengan hasil yang didapat pada harga sepatu dewasa \pm Rp. 210.000, dan harga sepatu anak-anak \pm 85.000.

3.2.3 Dokumentasi

Dokumentasi yang ada untuk mendukung penelitian ini berupa foto dengan pemilik toko *Sevendec Shoes*, serta data transaksi penjualan selama setahun.

3.3 Operasional Variabel

Operasional variabel merupakan sebuah cara yang digunakan sebagai alat bantu komunikasi antar penelitian agar lebih jelas lagi untuk sebagai arahan akan sebuah variabel tersebut. Variabelnya adalah :

3.3.1 Data Transaksi (Penjualan)

Data Transaksi merupakan jumlah dari banyaknya penjualan jenis sepatu selama sebulan dalam periode setahun terakhir. Dari data-data Transaksi (penjualan) ini, dapat ditentukan *item set* yang paling banyak terjual.

Tabel 3.2 Tabel Data Transaksi

Jenis Sepatu	Jan	Feb	Mar	Apr	Mei	Jun	Jul	Agst	Sep	Oktr	Nov	Des
Sneaker	35	20	10	0	11	15	20	12	12	0	22	0
Boots	11	6	0	14	13	0	0	9	0	8	0	12
Boots Anak	9	0	0	6	0	9	0	11	5	0	9	8
Flat Shoes	14	15	8	0	10	5	0	8	0	10	0	15
Sports	44	42	31	21	36	33	29	29	35	36	28	47
Sport Anak	0	9	0	11	8	0	11	9	0	12	8	0
Sling Back	5	0	0	3	0	3	0	0	2	0	0	1
Pantofel	9	0	5	0	0	11	7	0	0	8	0	7
Loafer	4	0	0	4	0	0	0	0	2	3	0	0
Slip On	0	0	0	6	0	0	1	2	0	4	0	0
Mules	3	0	0	1	0	0	0	0	0	2	0	0
Wedges	0	2	0	0	0	0	0	2	0	1	0	3
TST	6	0	0	4	0	4	0	0	3	0	5	6
High Heels	10	0	0	8	5	0	9	0	7	0	0	11

Sumber : Data Olahan Penelitian, 2019

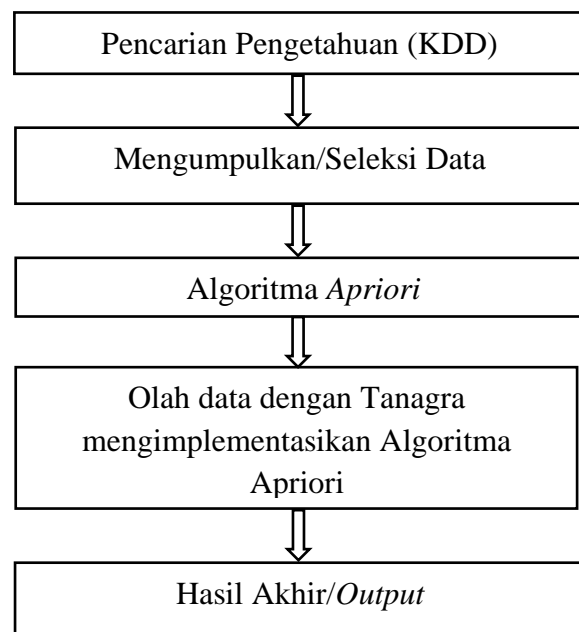
3.3.2 Jenis Sepatu

Jenis sepatu merupakan jenis atau macam sepatu yang ada di toko Sevendec Shoes yang terjual setiap bulannya. Jenis sepatu yang dimaksud seperti sneaker, flat shoes, boots, high heels, mules, sports, sling back shoes, pantofel, dll.

3.4 Metode Analisis dan Perancangan Sistem

3.4.1 Metode Analisis Perancangan Data Mining

Berikut perancangan pada data mining yang masing-masing akan dianalisis dimulai dari pencarian pengetahuan hingga *output* atau hasil akhir.



Gambar 3.3 Analisis Data Mining

Sumber : Olahan Data Penelitian, 2019

Berikut rangkuman dari gambar diatas akan dijelaskan :

1. Pencarian Pengetahuan (KDD)

Dengan melakukan penelitian langsung, toko *Sevendec Shoes* melakukan penjualan dan pendataan barang-barang dengan manual. Pendataan tersebut tidaklah akurat dikarenakan akan membuat penumpukan atau sisa barang yang banyak, serta menambah jumlah penyimpanan yang tinggi, serta dapat membuat kecewa para konsumen yang mencari jenis sepatu yang diinginkannya. Sehingga

pada tahap ini, dibutuhkan algoritma *Apriori* yang dapat mencari pola asosiasi antar satu barang dengan barang lainnya, juga dibutuhkan suatu sistem yang dapat membantu untuk mengetahui persediaan barang, serta pada saat penjualan tidak perlu melakukan *record* data secara manual.

2. Mengumpulkan/Seleksi Data

Dalam hal ini, peneliti mengumpulkan data dengan berupa transaksi penjualan barang di periode bulan Januari 2018.

3. Algoritma *Apriori*

Algoritma *Apriori* merupakan pola asosiasi atau hubungan antar satu atau lebih *item* dalam kumpulan data. artinya adalah penjualan sepatu dalam sehari dapat dijual berbagai jenis sepatu dan sesuai dengan konsumen atau pemakai mulai dari anak-anak hingga dewasa. Dalam hal ini, algoritma *apriori* bertindak untuk membangun pola hubungan atau asosiasi antara jenis sepatu satu dengan lainnya yang terjual dalam periode satu tahun.

4. Mengolah Data dengan Tanagra

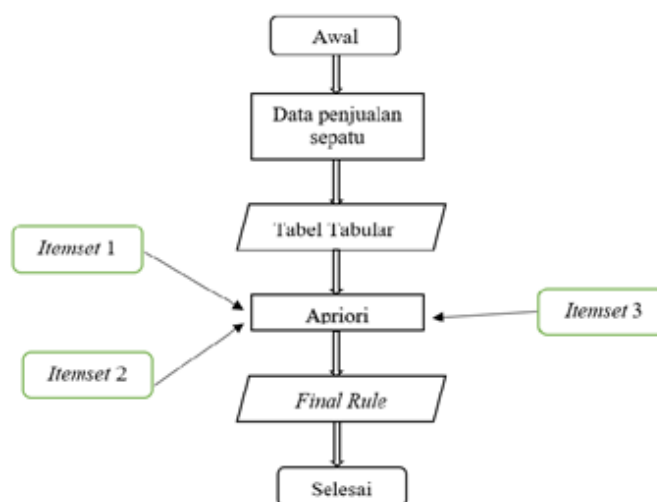
Data yang telah diolah dengan menggunakan algoritma *Apriori* akan dimasukkan terlebih dahulu ke perangkat lunak yang dipakai yaitu Tanagra. Berfungsi untuk mencocokkan kesamaan hasil data yang diolah dengan manual.

5. Hasil Akhir

Tahap yang terakhir adalah hasil, dengan menerapkan algoritma *apriori* akan diketahui jenis sepatu apa yang paling banyak terjual, sedang, kurang laku, atau bahkan tidak laku.

3.4.2 Perancangan *Data Mining*

Metode perancangan sistem dibuat kedalam penelitian dalam hal ini akan dijelaskan pada gambar dibawah ini :



Gambar 3.4 Rancangan Sistem
Sumber : Data Olahan Penelitian, 2019

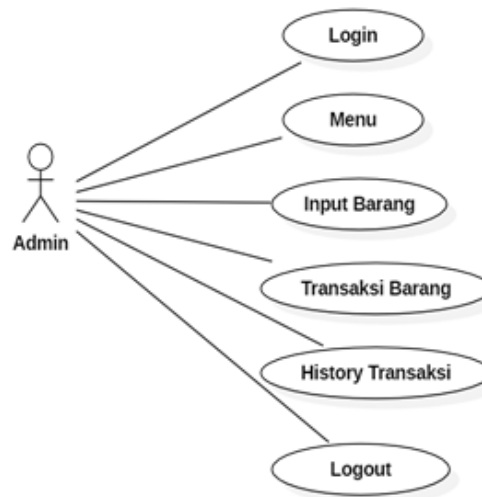
Pada gambar 3.4 di atas menjelaskan seluruh tahapan penelitian dari proses pengumpulan data.

3.4.3 Perancangan dan Desain Sistem dan Database

3.4.3.1 *Use Case Diagram*

Use Case dalam hal ini menggambarkan satu *actor* membuka sistem dan disebut sebagai *admin*. Cara kerja *Use Case* lebih menonjol dibagian fungsi sistem, yaitu dengan alur serta sistem. Pada sistem ini ada *login*, *menu*, *input* barang, transaksi barang, *history* transaksi, *logout*. Perlu diketahui Halaman Utama tidak

termasuk ke dalam diagram karena tampilannya sama dengan tampilan panel admin. Berikut akan dijelaskan *Use Case* di bawah ini



Gambar 3.5 Diagram *Use Case*
Sumber : Data Olahan Penelitian, 2019

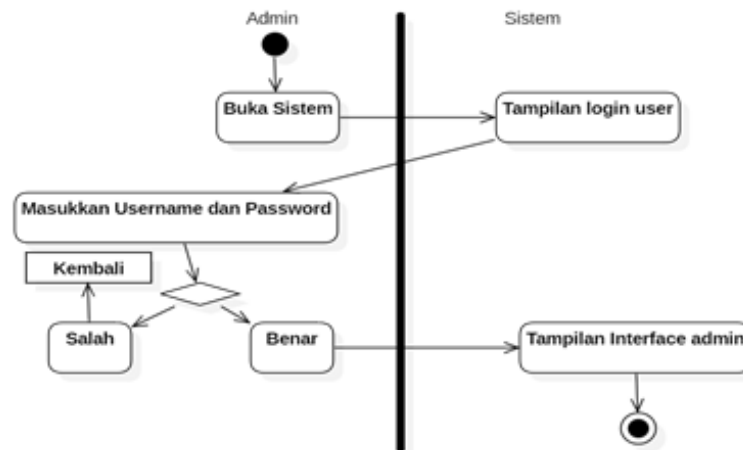
3.4.3.2 Activity Diagram

Activity Diagram menggambarkan aktivitas berupa kumpulan aktivitas.

Berikut penjelasan *Activity Diagram*.

1. *Login*

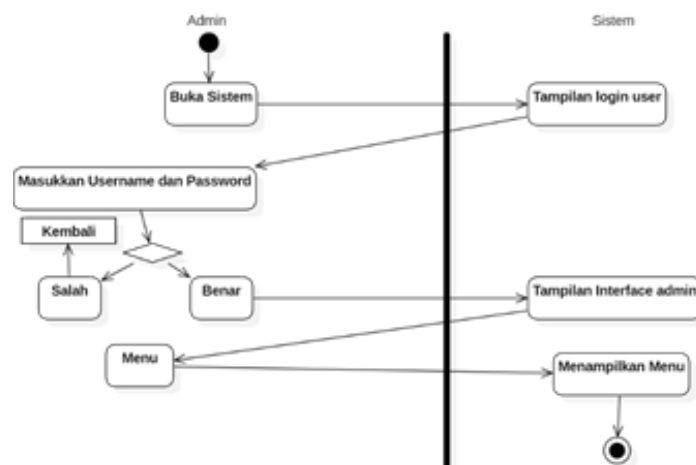
Dalam *Activity Diagram* ini, *admin* merupakan pemegang penuh atas akses sistem. *Admin* melakukan *login*, memasukkan *username* dan *password*, apabila salah maka akan kembali masukkan *username* dan *password*, jika benar akan menampilkan *panel admin*.



Gambar 3.6 *Login Diagram Activity*
Sumber : Data Olahan Penelitian, 2019

2. Menu

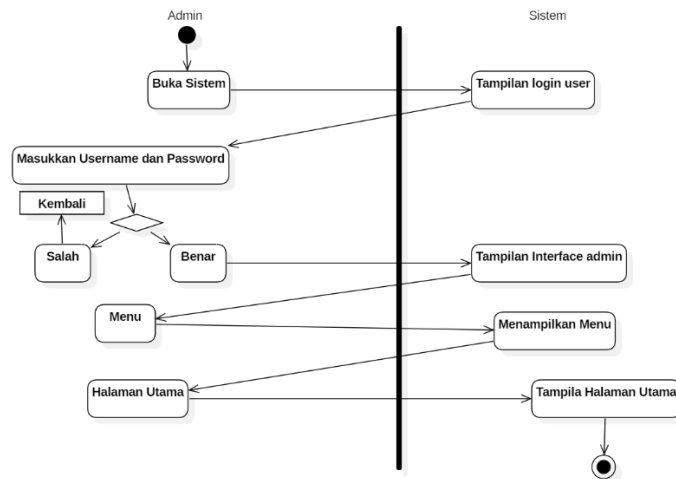
Admin merupakan pemegang penuh atas akses sistem. *Admin* melakukan *login*, *input username* dan *password*, tampilan *interface admin*, *menu*, tampilan *menu*. Berikut penjelasan *diagramnya*.



Gambar 3.7 *Menu Diagram Activity*
Sumber : Data Olahan Penelitian, 2019

3. Halaman Utama

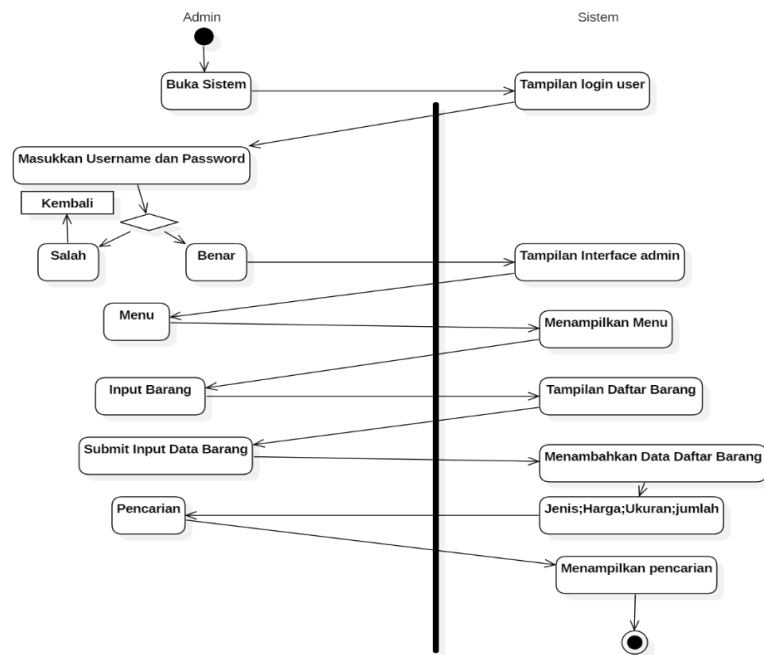
Admin merupakan pemegang penuh atas akses sistem. *Admin* melakukan *login*, *input username* dan *password*, tampilan *interface admin*, *menu*, tampilan *menu*, halaman utama, tampilan halaman utama.



Gambar 3.8 Halaman Utama Diagram *Activity*
Sumber : Data Olahan Penelitian, 2019

4. *Input* Barang

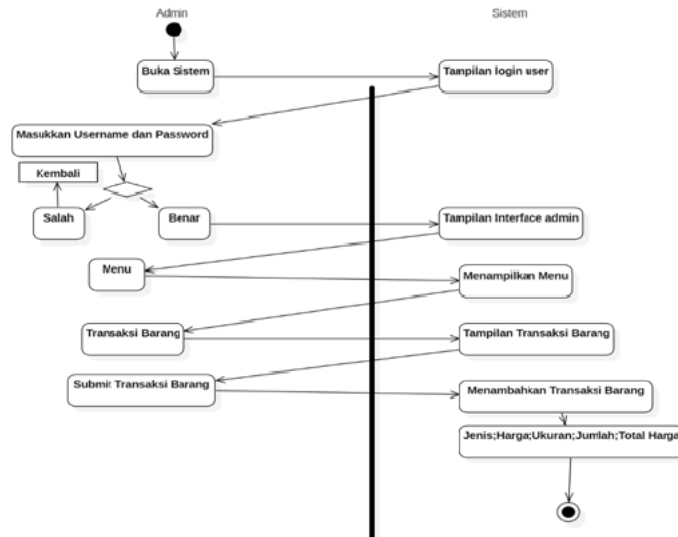
Admin merupakan pemegang penuh atas akses sistem. *Admin* melakukan *login*, memasukkan *username* dan *password*, tampilan *interface admin*, *menu*, tampilan *menu*, *input* barang, tampilan daftar barang, *submit input* data barang, menambahkan data daftar barang jenis, harga, ukuran, jumlah, pencarian, tampilan pencarian.



Gambar 3.9 *Input Barang Diagram Activity*
Sumber : Data Olahan Penelitian, 2019

5. Transaksi Barang

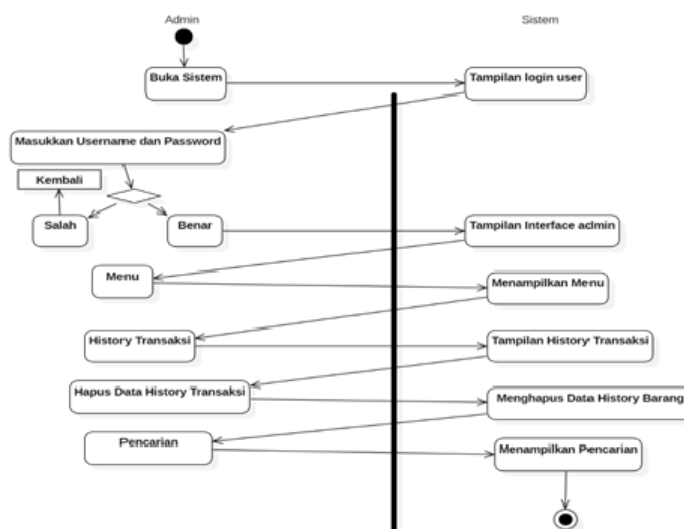
Admin merupakan pemegang penuh atas akses sistem. *Admin* melakukan *login*, *input username* dan *password*, apabila benar maka ke tampilan *interface admin*, *menu*, tampilan *menu*, transaksi barang, tampilan transaksi, *submit* atau tambahkan transaksi barang, menambahkan transaksi barang berupa Jenis, Harga, Ukuran, Jumlah, dan Total Barang.



Gambar 3.10 Transaksi Barang Diagram Activity
Sumber : Data Olahan Penelitian, 2019

6. *History* Transaksi

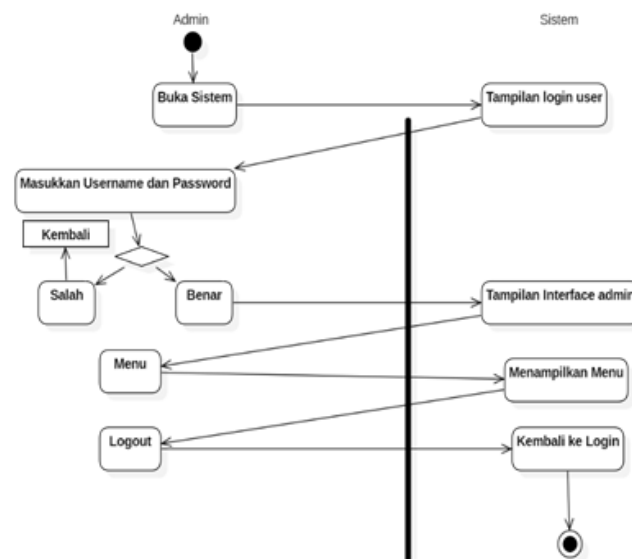
Admin merupakan pemegang penuh atas akses sistem. *Admin* melakukan *login*, *input username* dan *password*, apabila benar maka ke tampilan *interface admin*, *menu*, tampilan *menu*, transaksi barang, tampilan transaksi barang, lalu *submit* transaksi barang yang ingin ditambahkan, menambahkan transaksi barang berupa Jenis, Harga, Ukuran, Jumlah, dan Total Harga.



Gambar 3.11 *History* Barang Diagram Activity
Sumber : Data Olahan Penelitian, 2019

7. Logout

Admin merupakan pemegang penuh atas akses sistem. *Admin* melakukan *login*, *input username* dan *password*, apabila masuk benar maka ke tampilan *interface admin* dan *menu*, *logout*, dan kembali ke tampilan *login*.



Gambar 3.12 Logout Diagram Activity
Sumber : Data Olahan Penelitian, 2019

3.4.3.3 Sequece Diagram

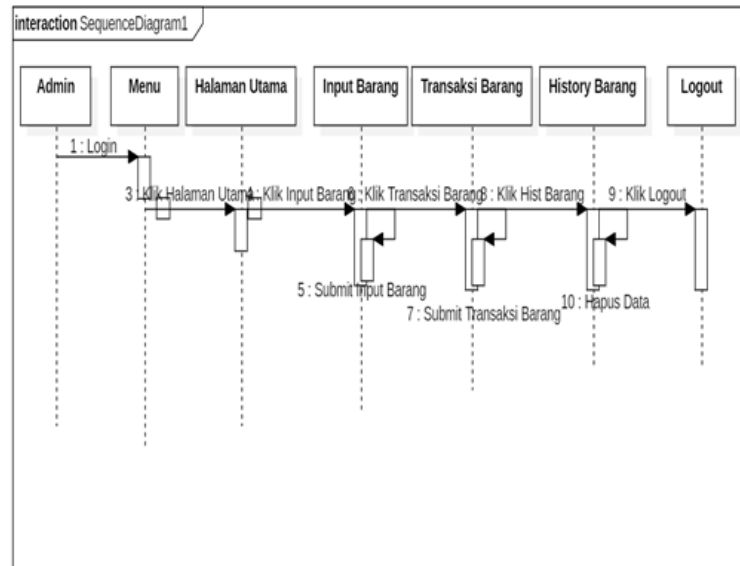
Diagram akan menunjukkan suatu gambaran sifat dan pola yang menggambarkan awal dari objek yang dibuat, serta menampilkan informasi yang diterima objek (Rosa & Shalahuddin, 2013).

1. Proses

Admin masuk ke dalam sistem dengan cara *login*. Lalu admin memiliki akses penuh untuk membuka/mengelola seluruh menu yang terdiri dari Halaman Utama, *Input Barang*, *Transaksi Barang*, *History Barang*, serta dapat *logout*. Pada menu *Input Barang* terdapat tombol *Submit* untuk menambahkan data Daftar Barang, *Transaksi Barang* juga memiliki tombol *Submit* untuk menambahkan transaksi yang

ada, serta pada *History* Barang terdapat tombol hapus data untuk menghapus data.

Berikut gambaran proses *Sequece* Diagram :

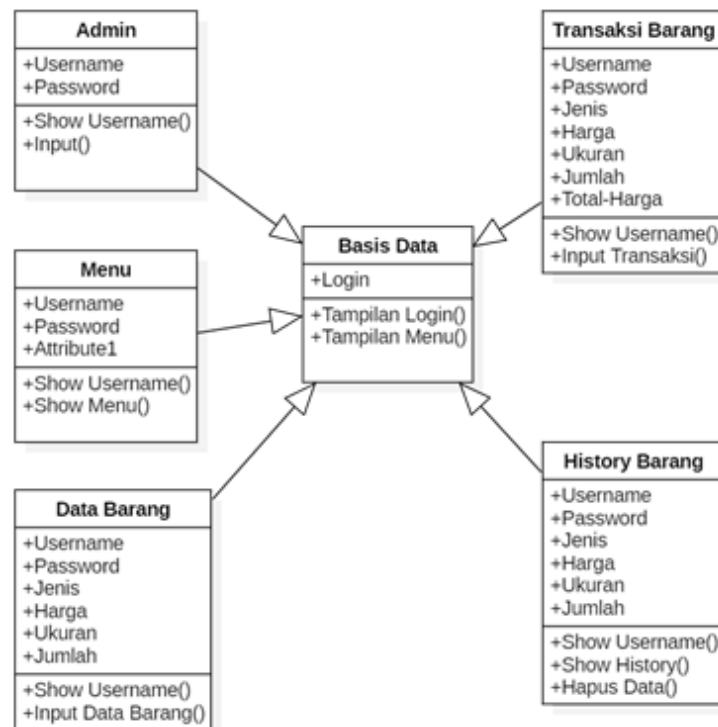


Gambar 3.13 *Sequence Diagram* Proses
Sumber : Data Olahan Penelitian, 2019

3.4.3.4 *Class Diagram*

Class Diagram disini menjelaskan bagian kelas utama yaitu, kelas *Admin*, kelas *Menu*, kelas *Input* Barang, kelas *Transaksi* Barang, dan kelas *History* Barang.

Diagram kelas pada sistem ini akan dijelaskan gambar di bawah :



Gambar 3.14 *Class Diagram*
Sumber : Data Olahan Penelitian, 2019

3.5 Lokasi dan Jadwal Penelitian

3.5.1 Lokasi Penelitian

Penelitian ini berada di toko sepatu Sevendec *Shoes* SP PLAZA 2, Lantai 2 Blok No. 05, Batu Aji, Batam.

3.5.2 Jadwal Penelitian

Penelitian ini memerlukan daftar kegiatan yang dilakukan dan berisi jadwal yang ada selama penelitian dilakukan. Berikut *list* jadwal penelitian yang dibuat dalam bentuk tabel :

Tabel 3.1 Jadwal Penelitian

No	Kegiatan	Kegiatan Penelitian																			
		September '19				Oktober '19				November '19				Desember '19				Januari '19			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Observasi Penelitian		■	■																	
2	Data Dikumpulkan				■	■	■														
3	Membuat BAB I							■	■												
4	Membuat BAB II								■	■											
5	Membuat BAB III										■	■	■								
6	Membuat BAB IV														■	■	■				
7	Membuat BAB V																		■	■	
8	Skripsi Dikumpulkan																				■

Sumber : Data Olahan Penelitian, 2019