

## BAB II

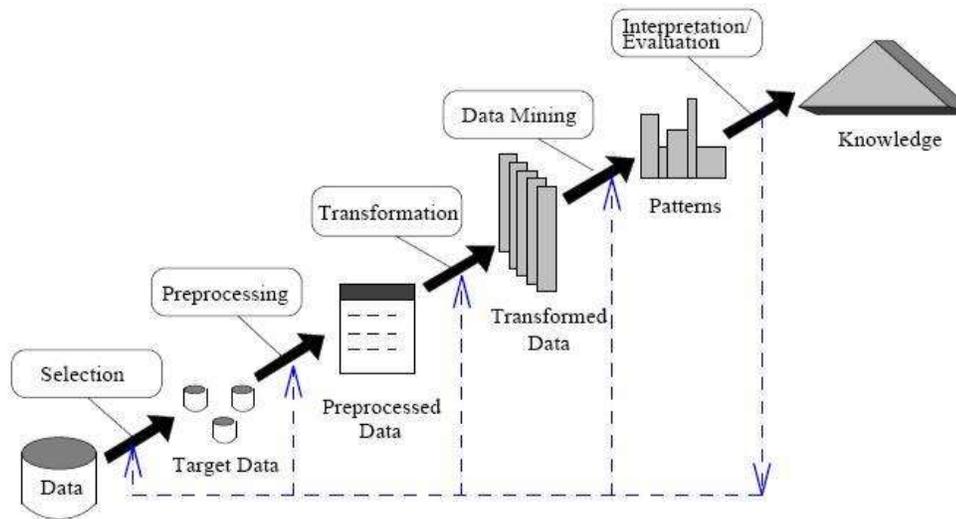
### KAJIAN PUSTAKA

#### 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 Vulandari, 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 Vlandari, 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 investasi serta sering butuh terhadap suatu metode agar hasil terakhir menjadi *valid* dan jelas.

## 2.3 Metode Data Mining

Pada data mining, 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 Vulandari, 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*

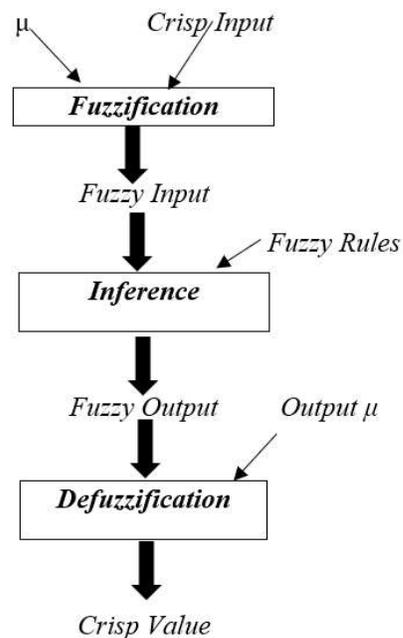
Metode ini mencari lalu menemukan beberapa  $k$  objek data ataupun pola terdekat dengan pola masukan, lalu memilih kelas dengan jumlah pola paling banyak diantara pola tersebut. Jika dicontohkan, metode ini melakukan klasifikasi pola dengan *voting*.

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 Vuldari, 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 mungkinya 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 nama <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 data-bbox="511 436 597 464"><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 berbisnis 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)

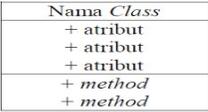
	Pemisahan organisasi-organisasi, serta menanggungjawab terhadap kegiatan saat terkini.
---	--

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
	Sama dengan konsep <i>interface</i> dalam pemrograman berorientasi objek.
	Hubungan antarkelas dengan artian pada umumnya.
	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

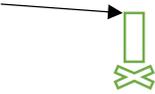
Diagram sekuen mendeskripsikan tingkah laku suatu *object* dalam *use case* dengan waktu hidup objek serta pesan yang terkirim dan masuk antara objek. Skenario pada *use case* dibutuhkan dalam membuat diagram sekuen ini dan seharusnya mengidentifikasi objek terkait di dalamnya.

**Tabel 2.4** Diagram *Sequence*

Simbol	Deskripsi
 <b>Nama aktor</b>	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 	Mendefinisikan objek yang berinteraksi pesan.

Sumber : Rosa, M. Shalahuddin, 2013

**Tabel 2.4** Diagram *Sequence* (Lanjutan)

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

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 ketetapan 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. Prisma 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*.

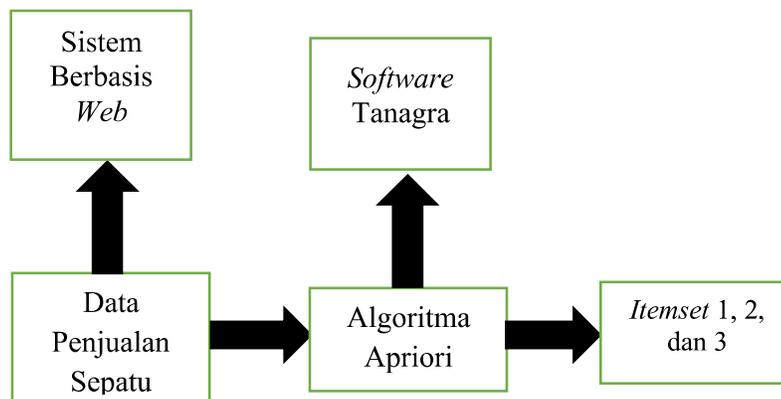
6. Penelitian Jiao Yabing yang berjudul “Research of An Improved Algorithm In Data Mining Association Rules” *explain that the classic algorithm of association rules, which enumerate all of the frequent item sets. When this algorithm encountered dense data due to the large number of long patterns emerge, 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 menganalisis 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