



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB II

LANDASAN TEORI

2.1 **Transportasi**

Transportasi adalah pemindahan barang dan manusia dari tempat asal ke tempat tujuan, dan salah satu jenis kegiatan yang menyangkut peningkatan kebutuhan manusia dengan mengubah letak geografis barang dan orang sehingga akan menimbulkan adanya transaksi (Utomo, 2010).

2.2 **Data Mining**

Data mining adalah proses logis untuk menemukan data yang berguna dari data dalam jumlah yang besar. Tujuan dari teknik ini adalah untuk menemukan pola yang sebelumnya tidak diketahui. Saat pola tersebut ditemukan, maka pola tersebut dapat digunakan untuk membuat keputusan untuk pengembangan suatu bisnis (Han, Kamber, & Pei, 2012).

Tiga langkah yang terlibat dalam *data mining* (Han, Kamber, & Pei, 2012):

1. *Exploration*

Langkah pertama dari eksplorasi data adalah dengan melakukan data cleansing. Variabel penting permasalahan dari data akan ditentukan.

2. *Pattern Identification*

Identifikasi dan memilih pola data agar membentuk informasi yang baik.

3. *Deployment*

Pola akan disebar untuk hasil yang diinginkan.

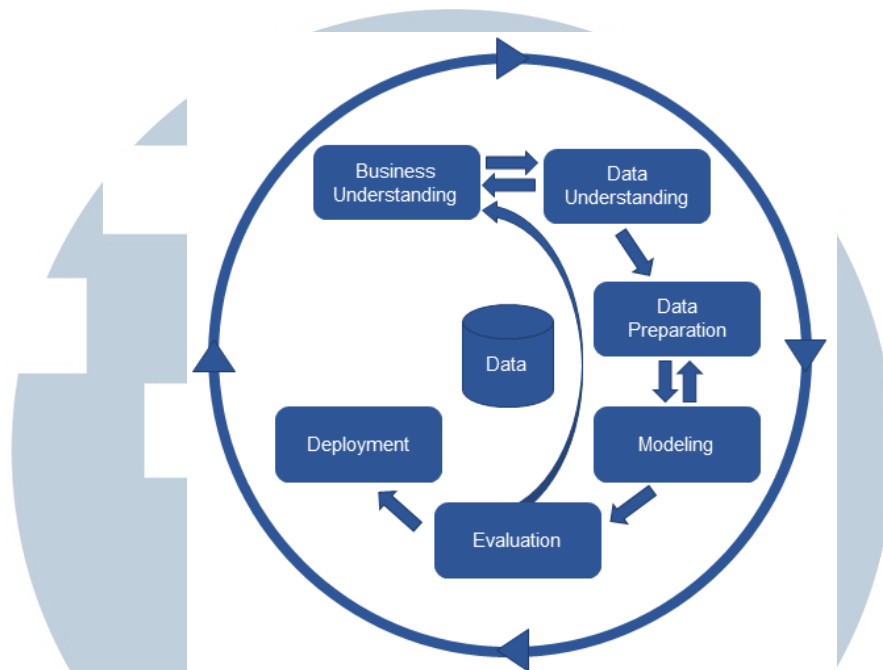
Menurut (Ian, Witten, Frank, Hall, & Pal, 2016) *Data mining* merupakan disiplin ilmu yang mempelajari metode untuk mengekstrak pengetahuan atau menemukan pola dari suatu data. *Data mining* adalah suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut (Singgih & Santoso, 2010).

2.3 **CRISP-DM**

CRISP-DM (*C*Ross-*I*ndustry *S*tandard *P*rocess for *D*ata *M*ining) untuk melakukan proses data mining terhadap data penjualan. CRISP-DM merupakan suatu konsorsium perusahaan yang didirikan oleh Komisi Eropa pada tahun 1996 dan telah ditetapkan sebagai proses standar dalam *data mining* yang dapat diaplikasikan di berbagai sektor industri. Berikut ini adalah enam tahap siklus hidup pengembangan *data mining* (Cirilio, 2017):

UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA



Gambar 2.1 CRISP-DM

Sumber : (Cirilio, 2017)

- a. *Business Understanding*: Tahap pertama adalah memahami tujuan dan kebutuhan dari sudut pandang bisnis, kemudian menterjemakan pengetahuan ini ke dalam pendefinisian masalah dalam *data mining*. Selanjutnya akan ditentukan rencana dan strategi untuk mencapai tujuan tersebut.
- b. *Data Understanding*: Tahap ini dimulai dengan pengumpulan data yang kemudian akan dilanjutkan dengan proses untuk mendapatkan pemahaman yang mendalam tentang data, mengidentifikasi masalah kualitas data, atau untuk mendeteksi adanya bagian yang menarik dari data yang dapat digunakan untuk hipotesa untuk informasi yang tersembunyi.

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

- c. *Data Preparation*: Tahap ini meliputi semua kegiatan untuk membangun dataset akhir (data yang akan diproses pada tahap pemodelan/*modeling*) dari data mentah. Tahap ini dapat diulang beberapa kali. Pada tahap ini juga mencakup pemilihan *table*, *record*, dan atribut-atribut data, termasuk proses pembersihan dan transformasi data untuk kemudian dijadikan masukan dalam tahap pemodelan (*modeling*).
- d. *Modeling*: Dalam tahap ini akan dilakukan pemilihan dan penerapan berbagai teknik pemodelan dan beberapa parameternya akan disesuaikan untuk mendapatkan nilai yang optimal. Secara khusus, ada beberapa teknik berbeda yang dapat diterapkan untuk masalah data mining yang sama. Di pihak lain ada teknik pemodelan yang membutuhkan format data khusus. Sehingga pada tahap ini masih memungkinkan kembali ke tahap sebelumnya.
- e. *Evaluation*: Pada tahap ini, model sudah terbentuk dan diharapkan memiliki kualitas baik jika dilihat dari sudut pandang analisa data. Pada tahap ini akan dilakukan evaluasi terhadap keefektifan dan kualitas model sebelum digunakan dan menentukan apakah model dapat mencapai tujuan yang ditetapkan pada fase awal (*Business Understanding*). Kunci dari tahap ini adalah menentukan apakah ada masalah bisnis yang belum dipertimbangkan. Di akhir dari tahap ini harus ditentukan penggunaan hasil proses *data mining*.

U
M
N
U
N
I
V
E
R
S
I
T
A
S
M
U
L
T
I
M
E
D
I
A
N
U
S
A
N
T
A
R
A

f. *Deployment*: Pada tahap ini, pengetahuan atau informasi yang telah diperoleh akan diatur dan dipresentasikan dalam bentuk khusus sehingga dapat digunakan oleh pengguna. Tahap *deployment* dapat berupa pembuatan laporan sederhana atau mengimplementasikan proses *data mining* yang berulang dalam perusahaan. Dalam banyak kasus, tahap *deployment* melibatkan konsumen, di samping analisis data, karena sangat penting bagi konsumen untuk memahami tindakan apa yang harus dilakukan untuk menggunakan model yang telah dibuat.

2.4 Database

Database adalah sekumpulan data tersebar yang berhubungan secara logis, dan penjelasan dari data ini dirancang untuk memenuhi kebutuhan informasi dari suatu organisasi (Connolly & Begg, 2010). Menurut (Gottschalk & Saether, 2010), *database* adalah sekumpulan data yang terorganisir untuk mendukung banyak aplikasi secara efisien dengan memusatkan data dan mengontrol *data redundant*.

2.5 Visualisasi Data

Definisi visualisasi adalah menggunakan teknologi komputer sebagai pendukung untuk melakukan penggambaran data visual yang interaktif untuk memperkuat pengamatan dan membantu eksplorasi data lebih dalam untuk membantu analisa data (Jeffrey, 2015).

Dengan menggunakan visualisasi, data yang ditampilkan dapat mempermudah peneliti untuk melihat data yang sulit dilihat dengan pemikiran sehingga peneliti bisa mengamati simulasi dan komputasi, juga memperkaya proses penemuan ilmiah dan mengembangkan pemahaman yang lebih dalam dan tak diduga, salah satu contohnya adalah dengan menampilkan data atau informasi dalam bentuk gambar, contoh: grafik, struktur pohon, pola, warna.

2.6 Power BI

Power BI merupakan kumpulan layanan perangkat lunak, aplikasi, dan konektor yang bekerja sama untuk mengubah sumber data yang tidak terkait menjadi saling berkaitan, ditampilkan dalam bentuk visual, dan interaktif (Power BI, Guided Learning, 2017).



Gambar 2.2 Logo Power BI

Sumber : Microsoft Docs

Power BI dapat dengan mudah menghubungkan data dalam bentuk Excel yang sederhana maupun data *warehouse* berbasis *cloud*. (Power BI, Guided Learning, 2017).

Power BI juga memungkinkan untuk melakukan visualisasi data dan penjelajahan data untuk menemukan hal penting dari data dan dapat dibagikan dengan siapa saja atau semua orang. (Power BI, Guided Learning, 2017).

Power BI terdiri dari desktop application yang disebut Power BI Desktop, online SaaS (Software as a Service) service yang disebut Power BI Service, dan aplikasi mobile Power BI yang dapat diunduh melalui Windows Phone, tablets, iOS, dan Android. Ketiga elemen tersebut, Desktop – service – Mobile dirancang untuk membiarkan orang untuk menciptakan, berbagi, mengkonsumsi wawasan bisnis sesuai dengan peran mereka dengan cara yang paling efektif. (Power BI, Guided Learning, 2017)

UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA