

# **BAB II**

## **TINJAUAN PUSTAKA**

### **2.1 Landasan Teori**

#### **2.1.1 Konsep Pemasaran**

Dalam merancang dan mengembangkan produk, baik yang berupa jasa maupun barang, tidak terlepas dari konsep pemasaran yang bertujuan memenuhi kebutuhan yang memuaskan pelanggan. Kepuasan pelanggan bisa, perilaku konsumen terhadap suatu produk dapat dideteksi dengan menarik kebutuhan pasar, menekan penetrasi pasar dengan teknologi baru, dan memodifikasi produk potensial untuk ditawarkan kepada konsumen.

Dalam konteks persaingan antara produsen, kita bisa mengidentifikasi faktor-faktor yang melatarbelakangi timbulnya kegiatan perancangan dan pengembangan produk. Faktor-faktor tersebut adalah sebagai berikut (Nasution, 2006:47) :

- a. Faktor eksternal, antara lain :
  1. Munculnya produk-produk sejenis dengan berbagai kelebihanannya.
  2. Munculnya produk-produk baru yang dapat menggantikan produk lama (produk substitusi).

3. Persegeran keinginan konsumen dan kebosanan terhadap produk-produk lama.
  4. Siklus hidup produk yang cenderung memendek pada masa modern ini.
- b. Faktor internal, antara lain :
1. Memperbaiki kinerja produk.
  2. Melakukan diversifikasi produk
  3. Mempertahankan segmen dan pangsa pasar baru.
  4. Memanfaatkan sumber daya manusia (karyawan, tenaga ahli) yang kemampuan semakin bertambah karena proses pembelajaran yang telah dialaminya.
  5. Menjaga kelangsungan hidup (keuntungan finansial) perusahaan.

### **2.1.2 Strategi Penjualan**

Dalam menentukan strategi penjualan sangat sulit dilakukan karena kita harus memahami segala kebutuhan dan keinginan dari seluruh konsumen yang ada sekarang ini untuk mengembangkan produk yang dapat memuaskan seluruh konsumen dalam ruang lingkup atau batasan sumber daya yang ada.

Variabel yang meliputi bauran penjualan antara lain ( Boone, 2006:7) :

#### **1. Produk**

Sebuah produk dirancang untuk memuaskan kebutuhan konsumen strategi produk meliputi sejumlah keputusan tentang kegunaan, kualitas, merk dagang, kemasan, desain, dan macam model lainnya.

## 2. Harga

Selain harga yang ditetapkan untuk sebuah produk yang di jual kepada konsumen, penentuan harga mencakup beberapa kebijakan manajemen mengenai diskon, harga, dan sebagainya.

## 3. Lokasi

Menempatkan produk berarti menyediakan produk pada tempat (pasar) yang tepat dan waktu yang tepat pula. Strategi penjualan distribusi produk meliputi sejumlah keputusan seperti lokasi dan daerah took.

## 4. Promosi

Promosi bermaksud untuk menginformasikan dan membujuk target konsumen dalam hal nilai dari prodek yang di jual. Sarana promosi utama adalah melalui iklan, penjualan secara besar-besaran.

## **2.2 Data Mining**

*Data Mining* adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam basis data. *Data Mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai basis data besar (Kusrini & Emha, 2009).

*Data mining* adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis (Fajar, 2013:3).

Kemajuan luar biasa yang terus berlanjut dalam bidang *data mining* didorong oleh beberapa faktor, antara lain (Larose, 2005) :

### 1. Pertumbuhan yang cepat dalam kumpulan data.

2. Penyimpanan data dalam data warehouse, sehingga seluruh perusahaan memiliki akses ke dalam basis data yang andal.
3. Adanya peningkatan akses data melalui navigasi web dan intranet.
4. Tekanan kompetisi bisnis untuk meningkatkan penguasaan pasar dalam globalisasi ekonomi.
5. Perkembangan teknologi perangkat lunak untuk *data mining*.
6. Perkembangan yang hebat dalam kemampuan komputasi dan pengembangan kapasitas media penyimpanan.

### **2.3 Pengelompokan *Data Mining***

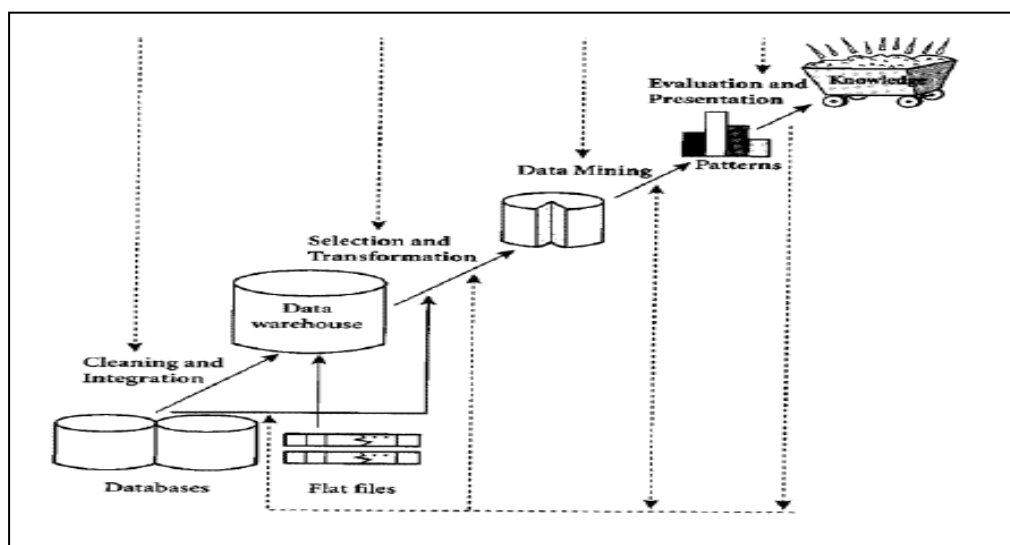
*Data mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu (Larose, 2005):

1. Deskripsi, terkadang peneliti dan analis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data.
2. Estimasi, estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numeric dari pada ke arah kategori. Model dibangun menggunakan record lengkap yang menyediakan nilai variabel target sebagai nilai prediksi.
3. Prediksi, prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang.
4. Klasifikasi, dalam klasifikasi, terdapat target variable kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

5. Pengklusteran, merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam kluster lain.
6. Asosiasi, asosisasi dalam *data mining* adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

## 2.4 Metode *Knowledge Discovery in Databases (KDD)*

Istilah *data mining* dan *knowledge discovery in databases (KDD)* seringkali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain. Dan salah satu tahapan dalam keseluruhan proses KDD adalah *data mining*. Proses KDD secara garis besar dapat dijelaskan sebagai berikut (Kusrini, 2009).



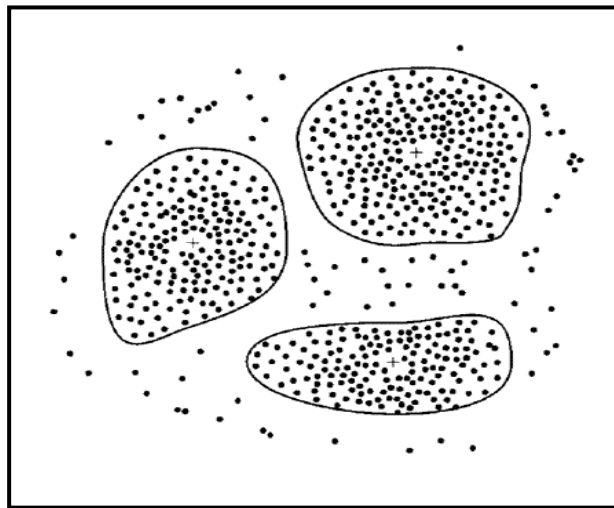
Gambar 2.1 Tahapan *Knowledge Discovery in Databases*

1. *Data Selection*, pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *KDD* dimulai. Data hasil seleksi yang akan digunakan untuk proses *data mining*, disimpan dalam suatu berkas, terpisah dari basis data operasional.
2. *Pre-processing/Cleaning*, sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus *KDD*. Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahn cetak.
3. *Transformation, coding* adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* dalam *KDD* merupakan proses kreatif dan sangat tergantung pada sejenis atau pola informasi yang akan dicari dalam basis data.
4. *Data Mining, data mining* adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam *data mining* sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *KDD* secara keseluruhan.
5. *Interpretation/Evaluation*, pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses *KDD* yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

## 2.5 Clustering

*Clustering* yaitu menemukan kumpulan objek hingga objek-objek dalam satu kelompok sama atau punya hubungan dengan yang lain dan berbeda atau tidak berhubungan dengan objek-objek dalam kelompok lain. Tujuan dari *clustering* adalah meminimalkan jarak di dalam *cluster* dan memaksimalkan jarak antar *cluster* (Fajar, 2013:123).

*Clustering* adalah memaksimalkan kesamaan antar anggota satu kelas dan meminimumkan kesamaan antar kelas/*cluster*. *Clustering* dapat dilakukan pada data yang memiliki beberapa atribut yang dipetakan sebagai ruang multidimensi. Ilustrasi dari clustering dapat dilihat di Gambar 2.2. dimana lokasi, dinyatakan dengan bidang dua dimensi, dari pelanggan suatu toko dapat dikelompokkan menjadi beberapa cluster dengan pusat cluster ditunjukkan oleh tanda positif (+).



Gambar 2.2. *Clustering*  
Sumber : Iko, 2003:4

Banyak metode *clustering* memerlukan fungsi jarak untuk mengukur kemiripan antar data, diperlukan juga metode untuk normalisasi bermacam atribut yang dimiliki data.

Beberapa kategori metode *clustering* yang banyak dikenal adalah metode partisi dimana pemakai harus menentukan jumlah  $k$  partisi yang diinginkan lalu setiap data dites untuk dimasukkan pada salah satu partisi (Iko, 2003:4).

## 2.6 Algoritma *K-Means*

Algoritma *K-Means* merupakan salah satu metode pengelompokan data *nonhierarki* (sekatan) yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik sama dimasukkan ke dalam satukelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan ke dalam kelompok yang lain. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diset dalam proses pengelompokan, pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok (Eko Prasetyo, 2012:178).

Algoritma *K-means* didasarkan pada minimasi fungsi objektif yang diformulasikan dalam persamaan :

$$J_m = \sum_{i=1}^N \sum_{j=1}^c (U_{ij})^m \|Z_i - C_j\|^2$$

Dengan :

$U_{ij}$  merupakan level keanggotaan dari  $Z_i$  dalam *Cluster*  $j$ <sup>2</sup>

$Z_i$  merupakan nilai data ke- $i$  dari  $d$ -dimensi data

$C_j$  merupakan nilai ke- $j$  dari  $d$ -dimensi *cluster center*

$M$  merupakan sembarang bilangan *real* lebih besar dari 1



Selanjutnya algoritma K-means disusun dengan langkah sebagai berikut :

Langkah 1 : Tentukan himpunan data  $Z$ . Tentukan jumlah *cluster* yang diharapkan  $1 < c < N$ , nilai pembobot  $m > 1$ , toleransi penghentian  $\epsilon > 0$ .

Langkah 2 : Inisialisasi matriks partisi secara acak,  $U(0) \in M^{fc}$ . Ulangi untuk  $l=3,4,5$

Langkah 3 : Hitung *cluster center (means)*.

$$V_i^{(l)} = \frac{\sum_{k=1}^N (U_{ik}^{(l-1)})^m Z_k}{\sum_{k=1}^N (U_{ik}^{(l-1)})^m}, 1 \leq i \leq c$$

Langkah 4 : Hitung jarak

$$D_{ikA}^2 = (Z_k - V_i^{(l)})^T A (Z_k - V_i^{(l)}), 1 \leq i \leq c, 1 \leq k \leq N$$

Langkah 5 : Perbaharui matriks partisi.

Untuk  $1 \leq K \leq N$

Jika  $D_{ikA} > 0$  untuk  $i = 1, 2, \dots, C$

$$u_{iK}^{(l)} = \frac{1}{\sum_{j=1}^c (D_{ikA} / D_{jKA})^{2/(m-1)}}$$

Atau dengan kata lain :

$$u_{iK}^{(l)} = 0 \text{ jika } D_{ikA} > 0 \text{ dan } u_{iK}^{(l)} \in [0,1] \text{ dengan } \sum_{i=1}^c u_{iK}^{(l)} = 1$$

Ulangi sampai  $\|U^{(l)} - U^{(l-1)}\| < \epsilon$ .

## 2.7 Weka

*Weka* adalah aplikasi *data mining open source* yang berbasis *java*. Aplikasi ini dikembangkan pada tahun 1994 dan pertama kali oleh sebuah universitas di selandia baru yang bernama universitas Waikato. aplikasi *weka* ini mulai menjadi aplikasi *data mining open source* yang sangat terkenal pada awal perkembangannya. Hal itu dikarenakan aplikasi *weka* memiliki kelebihan yang tidak dimiliki aplikasi *data mining* lainnya yaitu pada aplikasi *weka* terdapat banyak algoritma yang terdapat didalam aplikasi dan disertai juga *machine learning*, lalu juga dalam penggunaannya tidak terlalu rumit sehingga tidak menyulitkan penggunanya, dan ditambah dengan kelebihan lainnya bahwa algoritma-algoritma yang terdapat pada aplikasi *weka* selalu baru dan ter *update*, sehingga dengan beberapa kelebihan. Aplikasi *weka* tersebut digunakan banyak perusahaan dalam dunia bisnis untuk membantu dalam usaha bisnisnya, akademik pun juga tak mau ketinggalan untuk menggunakan aplikasi *weka*.

Aplikasi *weka* merupakan *software* yang terdiri dari koleksi algoritma *machine learning* yang dapat digunakan untuk melakukan generalisasi atau formulasi dari sekumpulan data *sampling*. Inti dari kekuatan pada aplikasi *weka* terletak pada algoritma yang makin lengkap dan canggih, namun walaupun begitu canggihnya aplikasi *weka* tersebut, letak keberhasilan *data mining* tetap ditentukan oleh manusia itu sendiri sebagai penggunanya/*user*. Keberhasilan *data mining* itu berdasarkan pada pengumpulan data yang berkualitas tinggi, penggunaan model dan algoritma yang tepat. Sehingga secanggih serta sehebat apapun aplikasi *data mining*, tanpa kemampuan sang penggunanya untuk menerapkannya maka tidak

akan menghasilkan data mining yang tepat dan bermanfaat(Feri & Dominikus, 2010;100).

## **2.8 Penelitian Sebelumnya**

Adapun penelitian sebelumnya yang digunakan untuk dapat dijadikan bahan pertimbangan dan di harapkan dapat membantu dalam pembuatan sistem yang baru.

### **2.8.1. penelitian Emha Taufiq Luthfi Dengan Judul Penerapan *Data Mining* Algoritma Asosiasi Untuk Meningkatkan Penjualan.**

Strategi penjualan yang tepat merupakan hal yang sangat penting dalam bisnis untuk dapat meningkatkan nilai penjualan. Banyak faktor yang dapat mempengaruhi tingkat penjualan suatu barang. Dalam tulisan ini dibahas penggunaan data *mining* algoritma asosiasi untuk menyusun sebuah sistem yang memiliki kemampuan melihat pola penjualan barang yang selanjutnya dapat digunakan untuk menyusun strategi penjualan baru. Algoritma asosiasi merupakan suatu bentuk algoritma dalam data *mining* yang memberikan informasi hubungan antar item data di database. Algoritma tersebut dapat dimanfaatkan secara luas dalam proses bisnis diantaranya dalam proses penjualan. Data *mining* algoritma asosiasi dapat membantu dalam proses penjualan dengan memberikan hubungan antar data penjualan yang dilakukan pelanggan sehingga akan didapat pola pembelian pelanggan. Pebisnis dapat memanfaatkan informasi tersebut untuk mengambil tindakan bisnis yang sesuai.

**2.8.2. penelitian Enur Irdiansyah Dengan Judul Penerapan *Data Mining* Pada Penjualan Produk Minuman Di Pt. Pepsi Cola Indobeverages Menggunakan Metode *Clustering*.**

Dalam dunia bisnis yang selalu dinamis dan penuh persaingan, para pelakunya harus senantiasa memikirkan cara-cara untuk terus survive dan jika mungkin mengembangkan skala bisnis mereka. PT. Pepsi Cola Indobeverages merupakan perusahaan yang bergerak dalam bidang industri minuman. Tidak hanya PT. Pepsi Cola Indobeverages, masih cukup banyak perusahaan-perusahaan yang bergerak dibidang serupa. Hal tersebut tentu saja menimbulkan persaingan bisnis antar perusahaan. Dalam rangka menghadapi persaingan bisnis dan meningkatkan pendapatan perusahaan, pihak terkait dalam perusahaan tersebut dituntut untuk dapat mengambil keputusan yang tepat dalam menentukan strategi pemasaran produk minuman yang akan dijualnya. Ketersediaan data yang melimpah, kebutuhan akan informasi (atau pengetahuan) sebagai pendukung pengambilan keputusan untuk membuat solusi bisnis, dan dukungan infrastruktur di bidang teknologi informasi merupakan cikal-bakal dari lahirnya teknologi *data mining*. *Data mining* dimaksudkan untuk memberikan solusi nyata bagi para pengambil keputusan di dunia bisnis, untuk mengembangkan bisnis mereka. Salah satu metode yang terdapat dalam *data mining* yang digunakan dalam penelitian ini adalah pengelompokan (*Clustering*) dimana metode tersebut mengidentifikasi objek yang memiliki kesamaan karakteristik tertentu, dan kemudian menggunakan karakteristik tersebut sebagai “vektor karakteristik” atau

“centroid”. Hasil dari penelitian ini adalah aplikasi yang dapat mempermudah menganalisis sejumlah data yang besar guna membantu memberikan informasi berharga sebagai gambaran dasar pengambilan keputusan perusahaan.