

## BAB III LANDASAN TEORI

### 3.1. *Business Intelligence*

*Business Intelligence* adalah sekumpulan teknik untuk mentransformasikan data mentah menjadi sebuah informasi yang berguna dan bermakna untuk tujuan analisis bisnis. Teknologi *Business Intelligence* digunakan untuk menangani data yang tidak terstruktur dalam jumlah yang sangat besar untuk membantu mengidentifikasi, mengembangkan, dan membuat strategi bisnis yang baru yang dapat dilakukan untuk tujuan organisasi. Terdapat banyak definisi tentang *Business Intelligence*, diantaranya:

- a. Menurut (Laudon dan Laudon, 2011), *Business Intelligence* adalah alat analisis yang digunakan untuk mengkonsolidasi data, menganalisis, menyimpan dan mengakses banyak data untuk membantu dalam pembuatan keputusan, seperti perangkat lunak untuk *query database* dan pelaporan alat untuk analisis data multidimensi, dan data mining.
- b. Menurut (Yulianton, 2008), *Business Intelligence* adalah semua penggalian modal bisnis untuk mendapatkan keuntungan dari data yang tersedia, baik yang tersebar pada sistem yang berbeda, maupun yang terintegrasi dalam tempat penyimpanan yang terpusat.

#### 3.1.1 Manfaat *Business Intelligence*

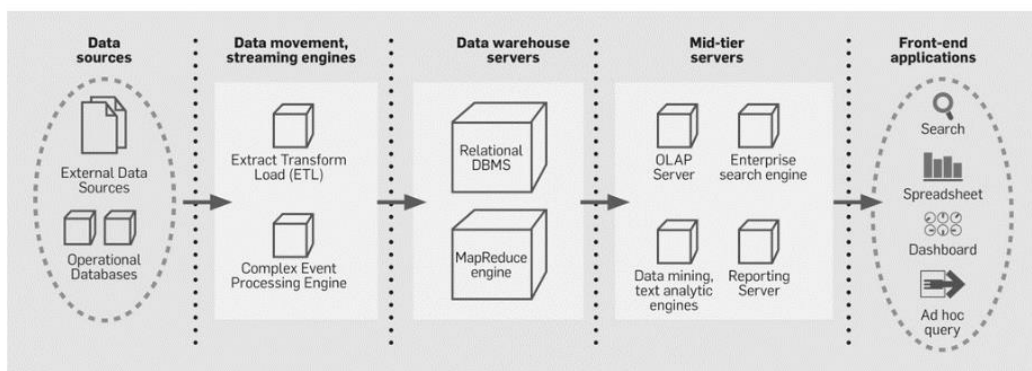
Ada banyak manfaat dan keuntungan yang diperoleh oleh perusahaan atau organisasi, dengan menerapkan *Business Intelligence* diantaranya menurut (Turban, Aronson dan Liang, 2005) adalah:

- a. Meningkatkan nilai data dan informasi organisasi. Dengan *Business Intelligence*, maka seluruh data dan informasi dapat diintegrasikan sehingga menghasilkan sebuah kesimpulan dari keadaan bisnis yang mudah di akses dan dimengerti, sehingga dapat membantu pengambilan keputusan.

- b. Memudahkan pengukuran kinerja organisasi di mana *Business Intelligence* dapat menunjukkan pencapaian *Key Performance Indicator* (KPI) suatu organisasi dengan mudah, cepat, dan tepat.
- c. Meningkatkan nilai investasi teknologi informasi yang sudah ada, di mana *Business Intelligence* tidak selalu menggantikan sistem informasi yang sudah ada, akan tetapi hanya menambahkan layanan pada sistem-sistem tersebut sehingga dapat menyajikan data dan informasi dengan lebih baik.
- d. Meningkatkan efisiensi biaya. *Business Intelligence* dapat meningkatkan efisiensi biaya karena dapat mempercepat seseorang dalam melakukan pekerjaan sehingga menghemat waktu dan mempermudah pemanfaatannya.

### 3.1.2 Arsitektur *Business Intelligence*

Menurut (Mishra dan Saini, 2016), arsitektur *business intelligence* terdiri dari *Data Sources*, *Data Movement and Streaming Engines*, *Data Warehouse*, *Servers Mid-tier servers*, dan *Front-end Application* yang kemudian di gambarkan sebagai berikut:



Gambar 3. 1 Arsitektur *Business Intelligence*, (Mishra dan Saini, 2016)

Adapun penjelasan masing-masing komponen dan tahapan arsitektur *Business Intelligence* adalah sebagai berikut:

a. *Data Sources*

Bagian ini berfungsi untuk pengumpulan data dan mengintegrasikan data yang disimpan dari berbagai sumber baik primer maupun sekunder, dapat mencakup dokumen terstruktur maupun yang tidak struktur seperti email dan data yang diterima dari penyedia eksternal, untuk kemudian akan di satukan dan diintegrasikan.

b. *Data Movement dan Streaming Engines*

*Data movement* dan *Streaming Engines* merupakan bagian proses menampilkan data dari setiap variasi data yang terintegrasi dan didapat dari berbagai sumber. Tugas *Data Movement* dan *Streaming Engines* ditangani sebuah alat berupa *Extract Transform Load (ETL)* yang membantu dalam menemukan masalah kualitas data dan memfasilitasi pemuatan data dengan jumlah yang besar kedalam warehouse. Sedangkan *The Complex Event Processing (CEP)* adalah sebuah alat yang pada tingkatan arsitektur fungsinya sama dengan ETL, tetapi ETL digunakan untuk proses *batch* data dan ketepatan waktu serta tidak kritis, sementara CEP digunakan untuk menangani real time atau hampir mendekati data yang real time. Dalam beberapa kasus, CEP digunakan untuk memvisualisasikan akses data *real time* untuk mendukung keputusan yang cepat.

c. *Data Warehouse Servers*

Data warehouse adalah tempat mengumpulkan subjek data yang berorientasi, terintegrasi, waktu yang bervariasi, dan non-volatile untuk mendukung proses pengambilan keputusan. Setelah data di ekstrak, terintegrasi dan diperiksa kualitas datanya kemudian dimuat dalam suatu repositori sentral atau yang disebut warehouse yang dikelola oleh satu atau lebih server. Data warehouse dirancang untuk mengenali setiap topik-topik yang menjadi perhatian sehingga para pengambil keputusan dapat menganalisis dengan mudah.

d. *Mid-tier servers*

Mid-tier servers menyediakan fungsi khusus untuk skenario *business intelligence* yang berbeda dan masuk dalam server *Online Analytical Processing (OLAP)*, *Enterprise Search Engines*, *Data Mining Engines*, dan *Reporting Servers*. Server OLAP yang efisien menyajikan model multidimensi untuk aplikasi *front end* atau langsung ke pengguna.

e. *Front-end Application*

*Front-end application* merupakan aplikasi yang digunakan langsung oleh pengguna bisnis dalam membuat keputusan. Sebagai contoh alat-alat dalam komponen *front-end application* adalah *enterprise portals for searching*, *spread sheets* dan aplikasi manajemen kinerja yang memvisualisasikan seperti dashboard atau alat-alat yang memungkinkan pengguna teknologi tersebut dapat mengakses dan melihat visualisasi model data.

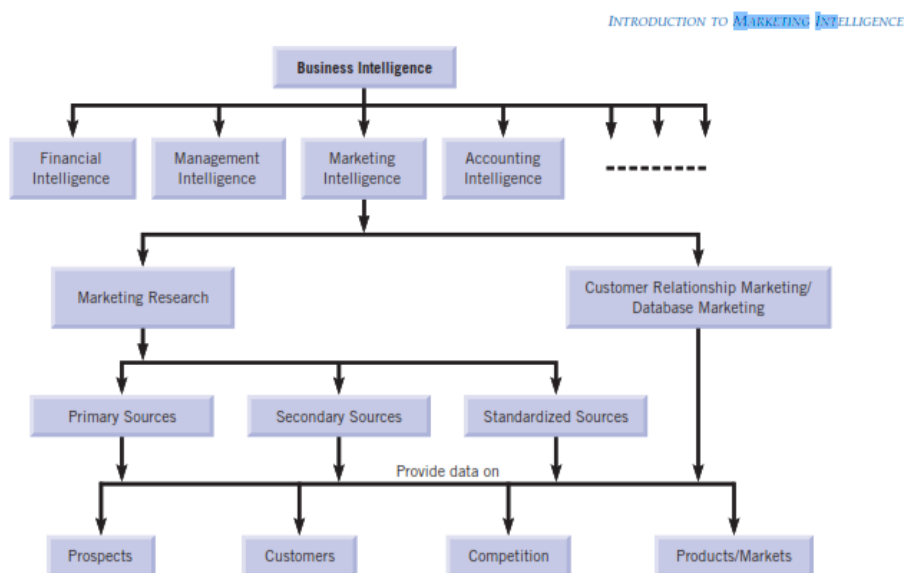
### 3.1.3 Penggunaan *Business Intelligence*

*Business Intelligence* dapat digunakan untuk meningkatkan nilai data dan informasi bagi organisasi. Dengan *Business Intelligence* maka dapat memudahkan pemantauan kinerja, meningkatkan nilai organisasi, menciptakan pegawai yang memiliki akses informasi yang baik, dan meningkatkan efisiensi biaya. Terkait dengan pemasaran dan promosi, beberapa kegunaan *Business Intelligence*, antara lain untuk :

- a. Analisa dalam perilaku konsumen, pola pembelian, dan trend penjualan
- b. Mengukur, melacak dan memprediksi penjualan dan kinerja keuangan
- c. Penganggaran dan perencanaan keuangan
- d. Mengetahui kinerja kegiatan pemasaran
- e. Optimalisasi proses dan kinerja organisasi promosi
- f. Analisa *Customer Relationship Management (CRM)*
- g. Analisa Resiko, nilai strategis, dan Media Sosial

### 3.2. *Marketing Intelligence*

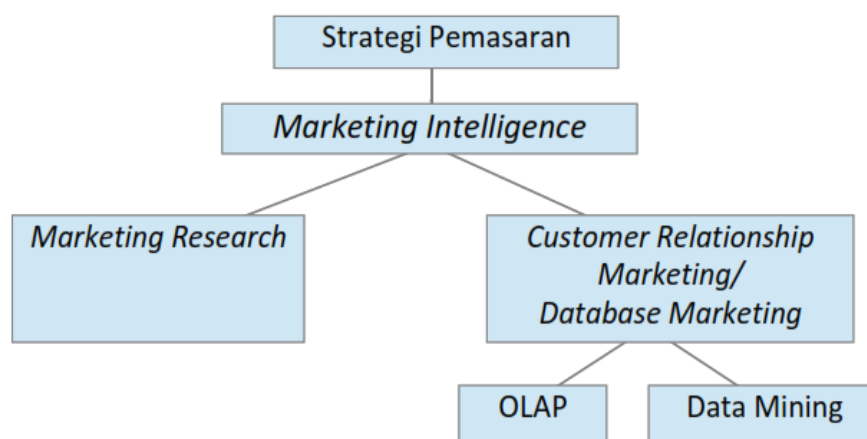
Menurut (Aaker *et al.*, 2012), *Marketing Intelligence* merupakan proses mengakusisi dan menganalisa informasi untuk memahami kondisi pasar terkait konsumen yang potensial, sikap dan perilaku pasar, dan perubahan dalam lingkungan bisnis, atau dengan kata lain *Marketing Intelligence* adalah kemampuan untuk mengerti, menganalisa, dan menilai lingkungan internal dan eksternal yang berhubungan dengan konsumen, pesaing, pasar, dan industri untuk mendukung proses pengambilan keputusan. Pada era globalisasi ini, *Marketing Intelligence* sangat dibutuhkan oleh perusahaan atau organisasi untuk mengetahui perilaku konsumen, para pesaing tanpa memata-matainya, keputusan manajemen memberikan target penjualan ke *sales*, mengidentifikasi produk baru, dan mengetahui ketidakpastian pasar pada masa depan sehingga manajemen dapat membuat perencanaan sebelumnya. *Marketing Intelligence* merupakan salah satu pendukung implementasi dari *Business Intelligence* selain *Financial Intelligence*, *Management Intelligence*, *Accounting Intelligence*, dan lain sebagainya.



Gambar 3. 2 *Marketing Intelligence* dalam *Business Intelligence*, (Aaker *et al.*, 2012)

Dari gambar diatas diketahui bahwa *Marketing Intelligence* terdiri atas dua bagian yaitu *Marketing Research* dan *Customer Relationship Marketing/Database*

*Marketing. Marketing Research* merupakan bagian *Marketing Intelligence* yang fokus pada proses perencanaan dari mulai menganalisa situasi pasar, mempersiapkan strategi, merancang program marketing, sampai dengan mengimplementasi strategi kepada target promosi, dengan menggunakan data dan informasi yang dimiliki, serta di dukung oleh data pendukung lainnya. Sedangkan *Customer Relationship Marketing/Database Marketing* lebih fokus pada pengolahan data marketing dalam sebuah database untuk kemudian digunakan sebagai pengambilan keputusan. Menurut (Guoxiang dan Zhiheng, 2013), *Database Marketing* adalah suatu teori baru yang menggabungkan teori teknologi informasi dengan teori pemasaran untuk melakukan proses pencarian pengetahuan baru guna mendukung pengambilan keputusan. Sedangkan menurut (Amborowati dan Suyanto, 2015), pada *Database Marketing* dilakukan proses pengumpulan data dan analisis informasi tentang konsumen dengan berbagai teknik analisa yang kemudian digunakan untuk mendukung pembuatan keputusan untuk membangun strategi pemasaran.



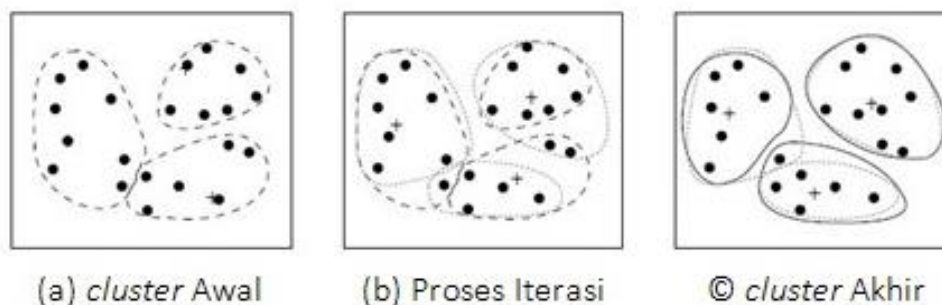
Gambar 3. 3 Dukungan *Marketing Intelligence* dalam strategi pemasaran  
(Amborowati dan Suyanto, 2015)

### 3.3. *K-Means Clustering*

*K-Means Clustering* adalah suatu metode pengelompokan data untuk melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Terdapat

dua jenis data *Clustering* yang sering dipergunakan dalam proses pengelompokan data yaitu *Hierarchical* dan *Non-Hierarchical*, di mana K-Means merupakan salah satu metode data *Clustering Non-Hierarchical* atau *Partitional Clustering*. K-Means adalah salah satu algoritma *Clustering* data yang bersifat *Unsupervised Learning*, yang berarti masukan dari algoritma ini menerima data tanpa label kelas. Tujuan dari metode *unsupervised learning* salah satunya adalah melakukan *Clustering*, yaitu mengelompokkan data-data dengan karakter mirip (Han, Pei dan Kamber, 2011).

Tujuan dari K-Means sendiri membagi suatu data dalam beberapa kluster (kelompok) sebanyak K, di mana jumlah K-nya ditentukan oleh Anda dan diwakili oleh Mean (Rata-rata). Mean dari setiap kluster diasumsikan sebagai ringkasan yang baik dari setiap observasi dari kluster tersebut. *K-Means Clustering* menentukan klasifikasi terhadap objek berdasarkan atribut kedalam K kluster. K adalah angka positif yang menyatakan jumlah grup/kluster/partisi terhadap objek. Pemartisian data dilakukan dengan mencari nilai jarak minimum antara data dan nilai centroid. Centroid adalah nilai rata-rata aritmetik dari sebuah bentuk objek dari seluruh titik dalam sebuah objek.



Gambar 3. 4 Proses *Clustering* Obyek Menggunakan metode K-Means, (Han, Pei dan Kamber, 2011)

#### **Langkah-langkah dalam Algoritma K-Means:**

1. Langkah 1: Tentukan berapa banyak kluster k dari dataset yang akan dibagi.
2. Langkah 2: Tetapkan secara acak data k menjadi pusat awal lokasi kluster.

3. Langkah 3: Untuk masing-masing data, temukan pusat kluster terdekat. Dengan demikian berarti masing-masing pusat kluster memiliki sebuah subset dari dataset, sehingga mewakili bagian dari dataset. Oleh karena itu, telah terbentuk kluster  $k$ :  $C_1, C_2, C_3, \dots, C_k$ .
4. Langkah 4: Untuk masing-masing kluster  $k$ , temukan pusat luasan kluster, dan perbarui lokasi dari masing-masing pusat kluster ke nilai baru dari pusat luasan.
5. Langkah 5: Ulangi langkah ke-3 dan ke-5 hingga data-data pada tiap kluster menjadi terpusat atau selesai.

### 3.4. *Confusion Matrix*

Menurut (Rahman *et al.*, 2017), *Confusion Matrix* adalah suatu metode yang digunakan untuk melakukan perhitungan akurasi pada konsep data mining atau Sistem Pendukung Keputusan. Pada *Confusion Matrix*, terdapat 4 (empat) istilah sebagai representasi hasil proses klasifikasi, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN).

<i>Correct Classification</i>	<i>Classified as</i>	
	<i>Predicted "+"</i>	<i>Predicted "-"</i>
<i>Actual "+"</i>	<i>True Positives</i>	<i>False Negatives</i>
<i>Actual "-"</i>	<i>False Positives</i>	<i>True Negatives</i>

Gambar 3. 5 Proses *Confusion Matrix*, (Rahman *et al.*, 2017)

Untuk memudahkan memahami *Confusion Matrix*, dapat dijelaskan dengan menggunakan contoh kasus sederhana untuk memprediksi seorang Siswa pintar atau tidak sebagai berikut:

#### ***True Positive (TP)***

Merupakan data positif yang diprediksi benar. Contohnya, Siswa pintar dan dari model yang dibuat memprediksi bahwa Siswa tersebut pintar.

#### ***True Negative (TN)***



Merupakan data negatif yang diprediksi benar. Contohnya, Siswa tidak pintar dan dari model yang dibuat memprediksi bahwa Siswa tersebut tidak pintar.

***False Positive (FP)***

Merupakan data negatif namun diprediksi sebagai data positif. Contohnya, Siswa tidak pintar tetapi dari model yang telah memprediksi Siswa tersebut pintar.

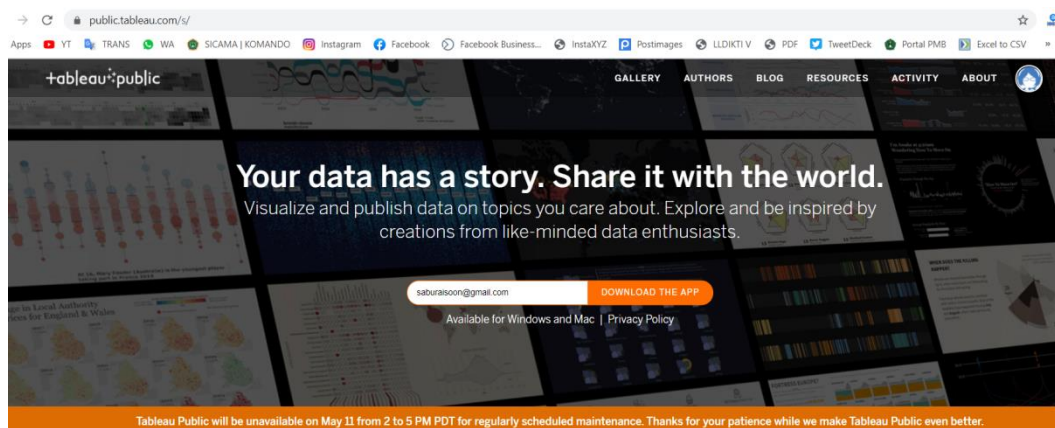
***False Negative (FN)***

Merupakan data positif namun diprediksi sebagai data negatif. Contohnya, Siswa pintar tetapi dari model yang dibuat memprediksi Siswa tersebut tidak pintar.

*Confusion Matrix* digunakan untuk memberi tahu seberapa baik model yang dibuat dengan memberikan informasi tentang TP, FP, TN, dan FN. Hal ini sangat berguna untuk mengetahui hasil klasifikasi sehingga dapat diketahui nilainya.

### **3.5. Tableau**

Tableau adalah sebuah aplikasi yang dirancang sebagai alat bantu dalam proses *Business Intellegent* untuk visualisasi data, sehingga memungkinkan para penggunanya melakukan analisis data mulai dari analisa struktur data, membuat grafik, sampai dengan membuat laporan yang sangat interaktif secara cepat. Tableau memiliki beberapa versi, yaitu *Tableau Desktop*, *Tableau Public*, *Tableau Online*, *Tableau Server*, dan *Tableau Reader*. Terdapat beberapa fungsi utama Tableau yang dapat membantu dalam proses *Business Intelligence*, diantaranya adalah mampu menerjemahkan dan memvisualisasikan data sehingga mudah di pahami. Tableau dapat digunakan secara online di webstite Tableau yang beralamat di <https://public.tableau.com> atau juga dengan menggunakan aplikasi dekstop yang dapat di unduh dari website tersebut dengan melakukan registrasi terlebih dahulu.



Gambar 3.6 Website Tableau (<https://public.tableau.com>)

Tableau saat ini banyak digunakan dalam pengolahan *Business Intelligence* karena memiliki beragam keunggulan di antaranya adalah:

- a. Pilihan visualisasi yang interaktif dan menarik di mana Tableau menyediakan berbagai pilihan visualisasi yang dinamis.
- b. *User friendly* di mana Tableau tidak mengharuskan penggunanya menguasai bahasa pemrograman. Selain itu Tableau sangat mudah digunakan, bahkan untuk pemula.
- c. Mampu mengolah data yang berasal dari berbagai sumber sekaligus, di mana Tableau mampu menggabungkan berbagai sumber data, big data, spreadsheet, cloud, dan berbagai tipe data lainnya.
- d. Memiliki *Dashboard Mobile Friendly* sehingga Tableau dapat kita gunakan mengolah data dengan hanya menggunakan perangkat mobile.
- e. Tableau terintegrasi dengan bahasa skrip, sehingga dapat digunakan untuk mengintegrasikan data dengan bahasa pemrograman yang biasa digunakan dalam data science, seperti bahasa R dan Python.