# BAB II

**LANDASAN TEORI**

## Konsep Data

### 1. Pengertian Data

Data menurut James O’Brian (2012: 37) “Dapat berbentuk termasuk data alfa numerik tradisional, yang terdiri dari angka dan huruf serta karakter lainnya yang menjelaskan transaksi bisnis dan kegiatan serta entitas lainnya.”

## Konsep Informasi

### 1.Pengertian Informasi

Menurut O’Brian (2012: 37) “Infromasi merupakan data yang telah dikonversi atau diolah ke dalam bentuk yang memiliki makna dan berguna bagi pihak secara spesifik. Sehingga data biasanyan dikenakan proses nilai tambah (pemrosesan data atau pemrosesan informasi). “

## Konsep Sitem Informasi

##  1.Pengertian Sistem Informasi

Menurut Satzinger (2012: 4), “Sistem informasi bisa diartikan sebagai seperangkat komponen yang saling berhubungan yang mengumpulkan (atau menerima), memproses, menyimpan, dan mendistribusikan informasi untuk mendukung pengambilan keputusan dan pengendalian dalam sebuah organisasi.“

## Konsep Data Mining

### Pengertian Data Mining

Penggalian data (bahasa Inggris: *Data Mining*) menurut Larose (2015:67) adalah “Ekstraksi [pola](https://id.wikipedia.org/wiki/Pola) yang menarik dari [data](https://id.wikipedia.org/wiki/Data) dalam jumlah besar. Suatu pola dikatakan menarik apabila pola tersebut tidak sepele, implisit, tidak diketahui sebelumnya, dan berguna. Pola yang disajikan haruslah mudah dipahami, berlaku untuk data yang akan diprediksi dengan derajat kepastian tertentu, berguna, dan baru.”

Penggalian data memiliki beberapa nama alternatif, meskipun definisi eksaknya berbeda, seperti KDD (*knowledge discovery in database*), analisis pola, arkeologi data, pemanenan informasi, dan intelegensia bisnis. Penggalian data diperlukan saat data yang tersedia terlalu banyak (misalnya data yang diperoleh dari [sistem basis data](https://id.wikipedia.org/w/index.php?title=Sistem_basis_data&action=edit&redlink=1) perusahaan*,*[*e-commerce*](https://id.wikipedia.org/wiki/E-commerce), data [saham](https://id.wikipedia.org/wiki/Saham), dan data [bioinformatika](https://id.wikipedia.org/wiki/Bioinformatika)), tetapi tidak tahu pola apa yang bisa didapatkan.

Data mining mempunyai fungsi yang penting untuk membantu mendapatkan informasi yang berguna serta meningkatkan pengetahuan bagi pengguna. Pada dasarnya, data mining mempunyai empat fungsi dasar yaitu:

1. Fungsi Prediksi *(prediction)*. Proses untuk menemukan pola dari data dengan menggunakan beberapa variabel untuk memprediksikan variabel lain yang tidak diketahui jenis atau nilainya.
2. Fungsi Deskripsi *(description)*. Proses untuk menemukan suatu karakteristik penting dari data dalam suatu basis data.
3. Fungsi Klasifikasi *(classification)*. Klasifikasi merupakan suatu proses untuk menemukan model atau fungsi untuk menggambarkan class atau konsep dari suatu data. Proses yang digunakan untuk mendeskripsikan data yang penting serta dapat meramalkan kecenderungan data pada masa depan.
4. Fungsi Asosiasi *(association).* Proses ini digunakan untuk menemukan suatu hubungan yang terdapat pada nilai atribut dari sekumpulan data

### Tahapan Data Mining

 Tahapan data mining menurut Menurut Jiawei Han (2012: 6), ada berberapa tahapan dalam data mining:

****

Sumber:Jiawei Han (2015:7)

##### Gambar 2.1

**Tahapan Data Mining**

1. Seleksi Data

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD ( *Knowledge Discovery In Data Base* ) dimulai. Data hasil seleksi yang digunakan untuk proses data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional.

1. *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.

1. Transformasi

adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data

1. Penggalian Data

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.

1. Interpretasi/evaluasi

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.

### Jenis Data Untuk Data Mining

Jenis dan kegunaan data manining Menurut Jiawei Han (2012:8), dapat dikategorikan menhadu berberapa jenis data yang bisa digunakan untuk data mining:

###### Database Data

Sistem database, juga disebut sistem manajemen basis data (DBMS), terdiri dari kumpulan data yang saling terkait dikenal sebagai database dan satu set program perangkat lunak untuk mengelola dan mengakses data. Program perangkat lunak menyediakan mekanisme untuk mendefinisikan struktur database dan penyimpanan data untuk menentukan dan mengelola akses data bersamaan, bersama, atau terdistribusi dan untuk memastikan konsistensi dan keamanan informasi yang tersimpan meskipun sistem mogok atau upaya akses yang tidak sah.

###### Gudang Data

Gudang informasi yang dikumpulkan dari berbagai sumber, disimpan di bawah skema terpadu, dan biasa berada di suatu tempat. Gudang data dibangun melalui proses pembersihan data, integrasi data, transformasi data, pemuatan data, dan penyegaran data berkala.

###### Data Transaksi

Secara umum, setiap *record* dalam database transaksional menyimpan transaksi, seperti pembelian pelanggan, pemesanan penerbangan, atau klik pengguna pada halaman web. Transaksi biasanya menyertakan nomor identitas transaksi unik (*transaction ID*) dan daftar item yang membentuk transaksi, seperti barang yang dibeli dalam transaksi. Database transaksional mungkin memiliki tabel tambahan, yang berisi informasi lain yang terkait dengan transaksi, seperti deskripsi item, informasi tentang wiraniaga atau cabang, dan sebagainya.

###### Data lainnya

Selain database relasional, data gudang data, dan data transaksi, masih banyak jenis data lain yang memiliki bentuk dan struktur serbaguna dan makna semantik yang agak berbeda. Data semacam itu dapat dilihat pada banyak aplikasi seperti data *time-related* atau *sequence* (misalnya catatan sejarah, data bursa saham, dan deret waktu dan data urutan biologis), aliran data (misalnya data surveilans dan sensor video, yang terus menerus ditransmisikan), data spasial (misalnya, peta), data desain teknik (misalnya, desain bangunan, komponen sistem, atau sirkuit terpadu), data *hypertext* dan multimedia (termasuk teks, gambar, video, dan data audio), grafik dan data jaringan (misalkan jejaring sosial dan informasi), dan web (gudang informasi besar yang didistribusikan secara luas yang tersedia di internet).

### Pelatihan Data(*Data Training*)

Menurut Han dan Kamber (2011 : 36), data mining adalah proses menemukan pola yang menarik dan pengetahuan dari data yang berjumlah besar. Lalu menurut Connolly dan Begg,(2010), Data mining adalah suatu proses ekstraksi atau penggalian data yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari database yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting. Dan menurut Vercellis (2009 : 77), data mining adalah aktivitas yang menggambarkan sebuah proses analisis yang terjadi secara iteratif pada *database* yang besar, dengan tujuan mengekstrak informasi dan *knowledg*e yang akurat dan berpotensial berguna untuk *knowledge* *workers* yang berhubungan dengan pengambilan keputusan dan pemecahan masalah.

### Proses Data Mining

Proses yang umumnya dilakukan oleh data mining antara lain: deskripsi, prediksi, estimasi, klasifikasi, pengelompokan dan asosiasi. Secara rinci proses data mining dijelaskan sebagai berikut (Larose, 2015):

#### Deskripsi

Deskripsi bertujuan untuk mengidentifikasi pola yang muncul secara berulang pada suatu data dan mengubah pola tersebut menjadi aturan dan kriteria yang dapat mudah dimengerti oleh para ahli pada domain aplikasinya. Aturan yang dihasilkan harus mudah dimengerti agar dapat dengan efektif meningkatkan tingkat pengetahuan (*knowledge*) pada sistem. Tugas deskriptif merupakan tugas data mining yang sering dibutuhkan pada teknik postprocessing untuk melakukan validasi dan menjelaskan hasil dari proses data mining. Postprocessing merupakan proses yang digunakan untuk memastikan hanya hasil yang valid dan berguna yang dapat digunakan oleh pihak yang berkepentingan.

#### Prediksi

Prediksi memiliki kemiripan dengan klasifikasi, akan tetapi data diklasifikasikan berdasarkan perilaku atau nilai yang diperkirakan pada masa yang akan datang. Contoh dari tugas prediksi misalnya untuk memprediksikan adanya pengurangan jumlah pelanggan dalam waktu dekat dan prediksi harga saham dalam tiga bulan yang akan datang.

#### Estimasi

Estimasi hampir sama dengan prediksi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan record lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi. Sebagai contoh, akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, berat badan, dan level sodium darah.

#### Klasifikasi

Klasifikasi merupakan proses menemukan sebuah model atau fungsi yang mendeskripsikan dan membedakan data ke dalam kelas-kelas. Klasifikasi melibatkan proses pemeriksaan karakteristik dari objek dan memasukkan objek ke dalam salah satu kelas yang sudah didefinisikan sebelumnya.

#### Pengelompokan

Pengelompokan merupakan pengelompokan data tanpa berdasarkan kelas data tertentu ke dalam kelas objek yang sama. Sebuah pengelompokan adalah kumpulan record yang memiliki kemiripan suatu dengan yang lainnya dan memiliki ketidakmiripan dengan *record* dalam pengelompokan lain. Tujuannya adalah untuk menghasilkan pengelompokan objek yang mirip satu sama lain dalam kelompok-kelompok. Semakin besar kemiripan objek dalam suatu cluster dan semakin besar perbedaan tiap cluster maka kualitas analisis cluster semakin baik.

#### Asosiasi

Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam suatu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja (*market basket analisys*). Tugas asosiasi berusaha untuk mengungkap aturan untuk mengukur hubungan antara dua atau lebih atribut.

## Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan (Decision Tree). Selain memiliki karakteristik sepert *ID3(Iterative Dichotomiser), C4.5* juga memiliki beberapa karakteristik yang berbeda yang merupakan perbaikan dari karakteristik *ID3*. Berikut ini beberapa karakteristik *C4.5* yang juga merupakan perbaikan terhadap *ID3*:

1. Dapat menangani atribut menarik.

2. Dapat menangani *missing value.*

3. Melakukan p*runing* untuk memperoleh model yang paling efisien.

4. Menggunakan kriteria *gain ratio* untuk menentukan jenis *split* yang terbaik.

Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut :

1. Pilih atribut sebagai akar.
2. Buat cabang untuk tiap-tiap nilai.
3. Bagi kasus dalam cabang
4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada.

## Kepuasan Konsumen

### Pengertian Kepuasan Konsumen

Kepuasan Konsumen, menurut Kotler dan Keller (2012:150), “Perasaan senang atau kecewa seorang yang muncul setelah membandingkan antara kinerja (hasil) produk yang di pikirkan terhadap kinerja (atau hasil) yang diharapkan. Jika kinerja dibawah harapan, pelanggan tidak puas. Jika kinerja memenuhi harapan, pelanggan puas, dan jika kinerja melebihi harapan, pelanggan amat puas.”

Beberapa hal yang dapat dilakukan dan diamati untuk mengukur kepuasan pelanggan adalah:

#### Analisa kehilangan pelanggan

Yaitu menganalisa apa saja penyebab pelanggan yang hilang, yang meninggalkan produk yang dimiliki perusahaan

#### *Survey* berkala

Perusahaan yang responsive mengukur kepuasan pelanggan dengan mengadakan *survey* berkala yaitu dengan cara mengirimkan daftar pertanyaan atau menelpon suatu kelompok acak dari pembeli mereka akhir – akhir ini untuk mengetahui perasaan mereka terhadap kinerja perusahaan.

#### Pembelanjaan perusahaan

Yaitu perusahaan yang bertindak sebagai pembeli dimana dengan demikian perusahaan dapat mengetahui apa saja kekurangan yang dimiliki.

Zeithmal dan Bitner (2009) mengemukakan bahwa kepuasan adalah konsep yang jauh lebih luas dari hanya sekedar penilaian kualitas layanan, namun juga dipengaruhi oleh faktor-faktor lain yang dapat dijelaskan sebagai berikut:

1. Kualitas layanan atau jasa, yaitu konsumen akan merasa puas apabila mereka mendapatkan layanan yang baik atau sesuai dengan yang diharapkan.

2. Kualitas produk, yaitu konsumen akan merasa puas apabila hasil mereka menunjukkan bahwa produk yang mereka gunakan berkualitas.

3. Harga, yaitu produk yang mempunyai kualitas yang sama tetapi menetapkan harga yang relatif murah akan memberikan nilai yang lebih

tinggi kepada konsumen.

4. Faktor situasi, yaitu keadaan atau kondisi yang dialami oleh konsumen.

5. Faktor pribadi dari konsumen, yaitu karakteristik konsumen yang mencakup kebutuhan pribadi.

## Transportasi

Menurut Miro (2005) ”Transportasi dapat diartikan usaha memindahkan, mengerakkan, mengangkut, atau mengalihkan suatu objek dari suatu tempat ke tempat lain, di mana di tempat lain ini objek tersebut lebih bermanfaat atau dapat berguna untuk tujuan-tujuan tertentu.”

Ada beberapa unsur pengangkutan, yaitu sebagai berikut:

1.  adanya sesuatu yang diangkut;

2.  tersedianya kendaraan sebagai alat angkut

3.  ada tempat yang dapat dilalui alat angkut.

Pengangkutan sebagai usaha memiliki ciri-ciri sebagai berikut:

1. Berdasarkan suatu perjanjian
2. Kegiatan ekonomi di bidang jasa
3. Berbentuk perusahaan
4. Menggunakan alat angkut mekanik

Transportasi manusia atau barang biasanya bukanlah merupakan tujuan akhir, oleh karena itu permintaan akan jasa transportasi dapat disebut sebagai permintaan turunan (derived demand) yang timbul akibat adanya permintaan akan komoditi atau jasa lainnya. Dengan demikian permintaan akan transportasi baru akan ada apabila terdapat factor- factor pendorongnya.

## Variabel Kepuasan

### 1. Variabel kepuasan dengan dimensi-dimensi yang terdiri dari:

#### Tangibles (Produk-produk fisik)

Fasilitas fisik, perlengkapan, peralatan, sarana komunikasi, dan lain-lain merupakan hal yang harus ada dalam proses jasa. Unsur ini merupakan sarana untuk mewujudkan layanan secara langsung sebagaimana yang disyaratkan dalam suatu transaksi barang maupun jasa. Sarana-sarana tersebut harus diwujudkan dalam bentuk yang menarik, informative, dan nyaman untuk konsumen sehingga nantinya sarana-sarana ini akan menunjang proses pemberian layanan kepada konsumen.

#### Reliability (Keandalan)

Merupakan kemampuan memberikan layanan yang dijanjikan dengan segera dan memuaskan .keandalan adalah kemampuan perusahaan untuk menampilkan layanan yang dijanjikan secara tepat dan konsisten. Keandalan dapat diartikan mengerjakan dengan benar sampai kurun waktu tertentu. Pemenuhan janji layanan yang tepat dan memuaskan meliputi ketepatan waktu dan kecakapan dalam menanggapi keluhan pelanggan serta pemberian layanan secara wajar dan akurat.

#### Responsiveness (Daya tanggap)

Daya tanggap yaitu kemampuan atau keinginan para karyawan untuk membantu dan memberikan jasa yang dibutuhkan pelanggan. Konsumen yang sedang melakukan pembelian atau pemilihan terhadap suatu produk mengharapkan keberadaannya diakui oleh penjual. Dalam hal ini, konsumen mengharapkan bahwa penjual mampu memberi informasi, pertolongan, dan respon yang positif. Perlu diperhatikan bahwa menjad responsif tidak sama dengan menjadi agresif.

#### Assurance (Jaminan)

Jaminan meliputi pengetahuan, kemampuan, keramahan, kesopanan, dan sifat dapat dipercaya. Hal ini bisa didapat dari kontak langsung antara penjual dan pembeli. Dengan adanya jaminan, pembeli tidak akan ragu dan terbebas dari bahaya dan risiko.

#### Emphaty (Empati)

Empati meliputi sikap kontak personel maupun perusahaan dalam memahami kebutuhan atau kesulitan pelanggan. Komunikasi yang baik dan penuh perhatian, akan memberi kemudahan dalam melakukan komunikasi atau hubungan. Empati juga bisa diwujudkan dalam rangkaian sikap supaya pelanggan merasa nyaman dan tidak salah tingkah.