Data Mining

Pertemuan 1



Bobot Penilaian



No	Jenis Kegiatan	Bobot (%)
1	Kehadiran dan partisipasi kuliah	15%
2	Tugas / Presentasi	20%
3	Ujian Tengah Semester	30%
4	Ujian Akhir Semester	35%
Jumlah		100%

A 85-100

A- 80-84,99

B+ 75-79,99

B 70-74,99

B- 65-69,99

C 55-64,99

D 45-54,99

E <44,99

Masalah eksplorasi data Ukuran data yang selalu mengalir pada jaringan komputer dan menghasilkan jumlah data yang sangat besar bahkan mencapai Terabytes atau peta bytes = 1 juta gigabytes.



Kebutuhan untuk berkompetisi dengan strategi yang tepat menjadi lebih tinggi.

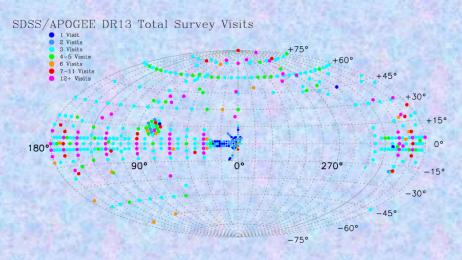
Beberapa kegiatan manusia yang memproduksi data:

- Mengakses world wide web / Log kunjungan Web.
- Riset science dan engineering / Akuisisi data dalam penelitianpenelitian. seperti; astronomi, kesehatan (menghasilkan rekam medis), dll.
- Transaksi penjualan, baik transaksi penjualan online maupun transaksi penjualan di supermarket.
- Transaksi perbankan dan kartu kredit.
- dll

Pertumbuhan Data

Astronomi

Menurut Sloan Digital Sky Survey, besar data yang di hasilkan setiap malam nya adalah 200 GB yang merupakan data fotometrik berupa bintang, galaksi, dll



(http://www.sdss.org, 2016)

Large Synoptic	Survey Telescope
----------------	------------------

- Chile, 2016
- Akan mendapatkan data 140TB setiap lima hari

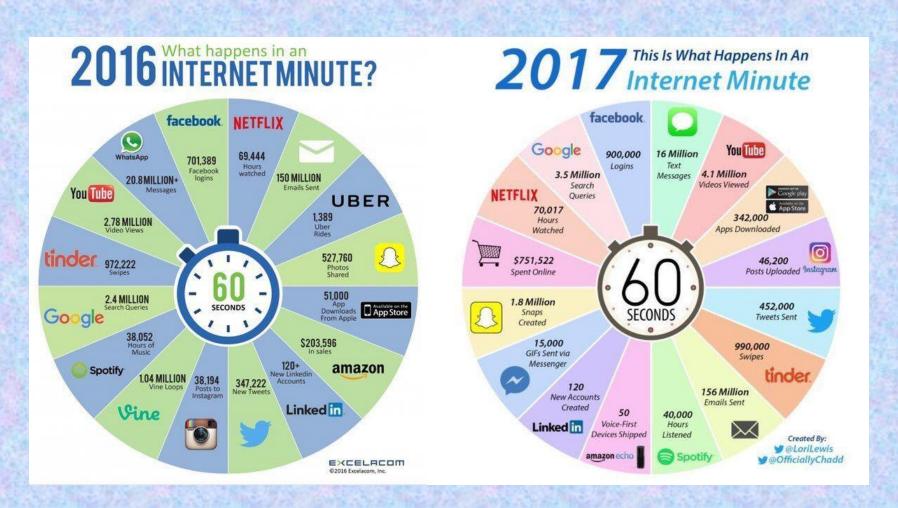
kilobyte (kB)	10 ³
megabyte (MB)	10 ⁶
gigabyte (GB)	10 ⁹
terabyte (TB)	10 ¹²
petabyte (PB)	10 ¹⁵
exabyte (EB)	10 ¹⁸
zettabyte (ZB)	10 ²¹
yottabyte (YB)	10 ²⁴

Datangnya Tsunami Data

- Mobile Electronics market
 - > 5Billion mobile phones in use in 2010
 - ➤ 150Million tablets was sold in 2012 (IDC)
 - 200Million is global notebooks shipments (Pengiriman) in 2012 (Digitimes Research)

- Web and Social Networks generates amount of data
 - Google processes 100 PB per day, 3 million servers
 - Facebook has 300 PB of user data per day
 - Youtube has 1000PB video storage
 - 235 TBs data collected by the US Library of Congress

Perubahan Culture dan Perilaku



(Insight, Big Data Trends for Media, 2017)

Mengapa Melakukan Data Mining??

Dilihat dari sudut pandang komersil

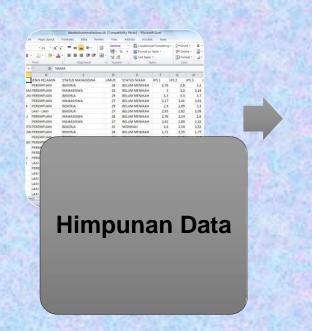
- ✓ Data over load
 - Web data, e-commerce, e-banking
 - Grocery stores
 - Bank/Credit Card Transaction
- ✓ Teknologi komputer menjadi lebih murah dan powerful
- ✓ Tekanan kompetisi semakin kuat
 - Dalam bisnis, dapat memberikan layanan yang lebih baik contoh: Customer Relationship Management

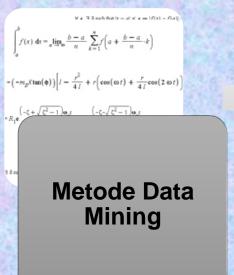
Mengapa Melakukan Data Mining??

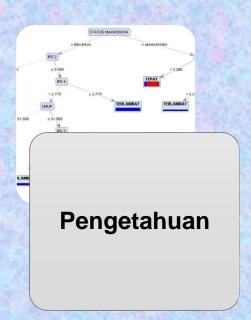
Dilihat dari sudut pandang keilmuan

- ✓ Data dikumpulkan dan disimpan dengan kecepatan tinggi (GB/hour)
 - Remote sensor pada satelite
 - Data spasial dalam GIS
 - Simulasi keilmuan
- ✓ Teknik tradisional tidak mampu melakukan analisis data yang berukuran besar
- ✓ Data Mining membantu ilmuan dalam
 - Klasifikasi dan segmentasi data
 - Pemodelan
 - Clustering

Apa itu Data Mining?







(romi@romisatriawahono.net)

Pengertian Data Mining

- Disiplin ilmu mengolah himpunan data yang sangat besar menjadi suatu pengetahuan, rumus, pola sehingga nanti dapat dimanfaatkan manusia untuk memprediksi kejadian kedepan dari kehidupan manusia.
- Melakukan ekstraksi untuk mendapatkan informasi penting yang sifatnya implisit dan sebelumnya tidak diketahui, dari suatu data (Witten et al., 2011)
- Ekstraksi informasi atau pola yang menarik (non-trivial, implicit, previously unknown dan potentially useful) dalam basis data berukuran besar (Han et al. 2011)