

Tugas KONSEP dan APLIKASI DATA MINING

Nama : Yolanda Menge

NIM : 17.51.0006

Prodi : Sistem Informasi

1. Sebutkan Karakteristik Big Data?

Karakteristik dari Big Data:

- a. Volume
Volume adalah: Besaran data yang harus dikelola berdasarkan super besar.
- b. Velocity
Velocity adalah: Berkaitan dengan kecepatan pemrosesan data yang harus mengimbangi pesatnya pertumbuhan jumlah data
- c. Variety
Variety adalah: Karakteristik sumber data yang sangat beragam, baik itu yang berasal dari basis data yang terstruktur maupun juga dari data-data yang tidak terstruktur.

2. Sebutkan Sumber Data Big Data?

Sumber Big Data Terbagi menjadi 2 yaitu:

- 1) Data Terstruktur
Kelompok data yang memiliki tipe data, format, dan struktur yang telah terdefinisi. Sumber datanya dapat berupa data transaksi, OLAP Data, tradisional RDBMS, File CSV, Spreadsheets.
- 2) Data Tidak Terstruktur
Kelompok data tekstual dengan format tidak menentu/tidak memiliki struktur melekat, sehingga untuk menjadikannya data terstruktur membutuhkan usaha, tools dan waktu yang lebih. Data ini dihasilkan oleh aplikasi-aplikasi internet, seperti data URL Log, media sosial, e-mail, blog, video dan audio.

3. Sebutkan 4 Tahapan Dalam Big Data?

Tahapan Dalam Big Data:

- a) Acquired
Berhubungan dengan sumber dan cara mendapatkan data
- b) Accessed

Berhubungan dengan daya akses data; data yang sudah dikumpulkan memerlukan data kelola, integrasi, storage dan computing agar dapat dikelola untuk tahap berikutnya. Perangkat untuk pemrosesan (processing tools) menggunakan Hadoop, Nvidia CUDA, Twitter Storm, dan GraphLab. Sedangkan untuk manajemen penyimpanan data (Storage tools) menggunakan Neo4J, Titan, dan HDFS (Hadoop Data File System).

c) Analytic

Berhubungan dengan informasi yang akan didapatkan, hasil pengelolaan data yang telah diproses. Analitik yang dilakukan dapat berupa descriptive (penggambaran data), diagnostic (mencari sebab akibat berdasarkan data), predictive (memprediksi kejadian dimasa depan) maupun prescriptive analytics (merekomendasikan pilihan dan implikasi dari setiap opsi). Tools untuk tahap analitik menggunakan MLPACK dan Mahout.

d) Application

Terkait visualisasi dan reporting hasil dari analitik. Tools untuk tahap ini menggunakan Rstudio.

4. Apa Manfaat Teknologi Big Data?

1) Pemanfaatan Big Data pada Sektor Bisnis

Perusahaan-perusahaan yang bergerak dalam sektor bisnis memiliki orientasi utama pada pencapaian margin laba setinggi mungkin (Profit Oriented). Berbagai informasi penting dapat dihasilkan dari Big Data yang dapat mendukung proses pengambilan keputusan bagi pimpinan perusahaan sebagai berikut:

- Mengetahui respon masyarakat terhadap produk-produk yang dikeluarkan melalui analisis sentimen di media sosial.
- Membantu perusahaan mengambil keputusan secara lebih tepat dan akurat berdasarkan data.
- Membantu meningkatkan citra perusahaan di mata pelanggan.
- Perencanaan usaha, dengan mengetahui perilaku pelanggan seperti pada perusahaan telekomunikasi dan perbankan.
- Mengetahui trend pasar dan keinginan konsumen.

2) Pemanfaatan Big Data Pada Sektor Layanan Publik

Perusahaan atau institusi yang berada pada sektor layanan publik lazimnya memiliki orientasi utama pada pencapaian kepuasan klien/pelanggan. Resource Big Data memberikan andil dengan menyajikan berbagai informasi berharga sbb:

- Mendapatkan feedback dan respon masyarakat sebagai dasar penyusunan kebijakan dan perbaikan pelayanan publik. Feedback tersebut dapat diperoleh dari sistem informasi layanan pemerintah maupun dari media sosial.

- Membuat layanan terpadu dengan segmen khusus sehingga layanan bisa lebih efektif dan efisien.
- Menemukan solusi atas permasalahan yang ada, berdasarkan data. Sebagai contoh: menganalisis informasi cuaca dan informasi pertanian terkait data tingkat kesuburan tanah, pemerintah dapat menetapkan atau menghimbau jenis varietas tanaman yang ditanam oleh petani pada daerah dan waktu tertentu.