**TUGAS**

Konsep & Aplikasi Data Mining



Oleh :

**Diahawa Fitaloka (17.51.0002)**

**PROGRAM STUDI S1- SISTEM INFORMASI**

**KEMENTRIAN RISET DAN TEKNOLOGI PENDIDIKAN TINGGI**

**SEKOLAH TINGGI MANAJEMEN INFORMATIKA KOMPUTER**

**PRADNYA PARAMITA**

**MALANG**

**2020**

Soal!

1. Sebutkan Karakteristik Big Data?
2. Sebutkan sumber data Big Data?
3. Sebutkan 4 tahapan dalam Big Data?
4. Apa mamfaat teknologi Big Data?

**Jawaban**

1. Karakteristik Big Data
2. Volume = Besaran data yang harus dikelola berukuran super besar
3. Velocity = Berkenaan dengan kecepatan pemrosesan data yang harus mengimbangi pesatnya pertumbuhan jumlah data
4. Veriety = Karakteristik sumber data yang sangat beragam,baik itu yang berasal dari basis data yang terstruktur maupun juga dari data-data yang tidak terstruktur.
5. Sumber data Big Data

Ada banyak sumber data big data,menurut professor munchan sumber big data ada 9 sumber big data yang paling sering digunakan:

1. Machine Data

Data yang diproduksi oleh mesin contohnya : *command,*hitungan,dan informasi yang *ditampilkan melalui manufacturing automation system.*

1. Knowledge

Misalnya : blogspot atau dokumentasi.

Dalam kehidupan sehari-hari kita sering menggunakan search engine. Proses pengolahan data yang dilakukan oleh mesin pencari tersebut tentunya memerlukan teknologi **big data.**

1. Sensor

Contoh sensor diantaranya adalah CCTV,GPS,Pengukur langkah kaki,detector panas,dll.

Misalnya, kacamata dengan kamera rahasia atau pen recorder. Karena itu sensor akan menjadi salah satu faktor utama dalam perkembangan big data beberapa tahun kedepan.

1. Ilmu Pengetahuan & Kesehatan

Data yang diproduksi oleh instrumen saintifik dan peralatan medis. Misalnya dari jaraingan pengawas cuaca otomatis,satelit,atau alat kesehatan seperti pemindai jaringan kanker.

1. Media

Data bisa memeiliki nilai artistic,entertainment,social,atau informative,dari konten-konten yang yang dibagikan oleh orang-orang misalnya gambar,video,tulisan dsb.

1. User Interface

Contohnya Aplikasi media sosial atau aplikasi lainnya yang bisa memperoduksi begitu banyak input data.

1. Transaksi

Transaksi komersil seperti pembelian melalui e-commerce atau bursa saham.

1. Komunikasi

Contohnya email,chat,dan komunikasi yang kita lakukan di media social.

1. Derived Data

Data yang dihitung atau di interpretasikan dari data lain. Contohnya metric marketing seperti brand engagement yang dihtung menggunakan data dari media social.

1. 4 Tahapan Pengelolaan Big Data
2. Acquired = Berhubungan dengan sumber dan cara mendapatkan data.
3. Accessed = Berhubungan dengan daya akses data, data yang sudah dikumpulkan memerlukan tata 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 menajemen penyimpanan data (storage tools) menggunakan Neo4j, Titan, dan HDFS (Hadop Data File System).
4. Analytic = Berhubungan dengan infromasi yang akan didapatkan, hasil pengelolaan data yang telah di proses. Analitik yang dilakukan dapat berupa descriptive   
   (penggambaran data), diagnostic (mencari sebab akibat berdasar data), predictive(memprediksi kejadian dimasa depan) maupun prescriptive analytics (merekomendasikan pilihan dan implikasi dari setiap opsi). Tools untuk tahap analiytic menggunakan MLPACK dan Mahout.
5. Application = Terkait visualisasi dan reporting hasil dari analitik. Tools untuk tahap ini menggunakan RStudio.
6. Mamfaat Teknologi Big Data.

Dalam kehidupan sehari-hari, kita sudah bisa melihat sebuah big data yang dapat digunakan oleh perusahaan-perusahaan untuk menganalisis data tersebut. Berikut ini adalah contoh pemanfaatan big data yang bisa digunakan dalam berbagai bidang :

* Pemamfaatan dalam bidang Bisnis Manufaktur

Perusahan manufaktur dapat memantau kondisi peralatan setiap saat (real-time), sehingga dapat memperkirakan waktu terbaik untuk mengganti peralatan. Karena mengganti terlalu cepat akan merugikan/buang-buang uang atau kalau terlambat akan menyebabkan produksi terganggu karena kerusakan peralatan.

Perusahaan manufaktur juga bisa memantau produk yang baru launching melalui social sedia untuk mengetahui apakah ada isu *after-sales* sehingga dapat mencegah kegagalan garansi yang menyebabkan publikasi besar yang dapat merusak citra produk dan perusahaan.

* Pemanfaatan dalam bidang Pertanian

Pada bagian pertanian ini, peneliti bisa mengambil sebuah data mulai dari kondisi tanah, cuaca, memperhatikan pertumbuhan sebuah tanaman dalam jangka waktu tertentu. Data tersebut nantinya akan dikumpulkan dan dianalisis. Hal ini tentunya akan membantu para petani untuk menentukan kondisi seperti apa yang akan membuat tanaman menjadi lebih subur dan menghasilkan hasil yang lebih baik.

* Pemanfaatan dalam bidang Perpajakan

Penggunaannya pada bidang perpajakan di Indonesia ini sebenarnya masih dalam tahap pengembangan. Pada sistem pajak online, data yang didapat terbilang sangat detail. Data tersebut kemudian divisualisasikan seperti silsilah keluarga, jenis dan barang kekayaan apa saja yang dimiliki, serta jenis pajak dan status apakah sudah melakukan pembayaran atau belum dari setiap orang yang memiliki wajib pajak di Indonesia.

Dengan data-data tersebutlah pemerintah bisa memanfaatkan untuk mengoptimalisasi pendapatan negara dengan pajak.

* Pemanfaatan dalam bidang Kesehatan

Pada bidang kesehatan, data-data yang ada dari setiap pasien yang sakit mulai dari klinik, puskesmas sampai rumah sakit akan dikumpulkan, diolah dan akan dianalisis. Dengan begitu, semua informasi kesehatan penduduk Indonesia akan menjadi terpusat. Hal ini akan memudahkan pemerintah untuk dapat melihat tingkat kesehatan seluruh penduduk Indonesia juga pemerintah dapat melakukan prediksi terhadap penyakit yang ada di Indonesia.