

LAPORAN
RENCANA TUGAS MANDIRI (RTM) Ke-1
MATA KULIAH BIG DATA
“MERINGKAS TOPIK KONSEP BIG DATA
DAN ANALITIK DATA”



DISUSUN OLEH:

Citra Amelia Intan Permadani (21083010004)

DOSEN PENGAMPU:

Tresna Maulana Fahrudin S.ST., M.T. (NIP. 199305012022031007)

PROGRAM STUDI SAINS DATA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAWA TIMUR
2022

a. Definisi, Karakteristik, dan Arsitektur Big Data

Big Data adalah istilah untuk menggambarkan suatu kumpulan data yang memiliki ukuran/volume yang sangat besar, kompleks, dan terus bertambah.

Dikatakan sebagai Big Data apabila data memenuhi karakteristik Big Data sebagai berikut:

- Volume → mengacu pada jumlah data yang sangat besar
- Velocity → mengenai kecepatan data dapat dihasilkan, diproses, dan dianalisis dalam memenuhi kebutuhan perusahaan.
- Variety → variasi format data yang berbeda-beda, yakni data terstruktur (data berbentuk tabular sehingga dapat digunakan langsung), semi-struktur (JSON dimana perlu diolah sebelum digunakan), dan data tidak terstruktur (data berbentuk teks, audio, video, dan gambar yang perlu dianalisis, diolah sehingga baru bisa digunakan).
- Veracity → keakuratan sebuah data
- Value → adanya value/nilai-nilai dalam big data memudahkan pengguna mengakses informasi secara cepat dan mengambil keputusan berdasarkan data yang tersedia.

Dalam proses pengolahan Big Data memiliki arsitektur yang jelas, yakni:

- Data Source → sumber data untuk Big Data
- Data Storage → tools untuk penyimpanan data hasil pemrosesan baik realtime maupun non-realtime, misalnya Hadoop, HDFS
- Data Processing → tools untuk memproses pengolahan Big Data, misalnya MapReduce
- Data Access → tools untuk melihat dan mengakses Big Data yang nantinya dapat dimasukkan ke data storage, misalnya Pig dan Hive
- Management → tools sebagai distribusi data ke sistem lain, misalnya Apache Zookeeper

b. Tipe Analitik, Mode Pemrosesan Analitik, dan Implementasi Analitik Big Data

Tipe Big Data Analitik:

- Big Data Analitik Preskriptif → menggunakan proses data analisis prediktif dan deskriptif dalam memberikan insight untuk dapat memperoleh hasil yang sesuai dengan apa yang telah diprediksikan.
- Big Data Analitik Prediktif → membuat prediksi tentang hasil di masa depan yang andal dan lebih akurat berdasarkan wawasan dari data dengan menggunakan alat dan model prediktif canggih seperti AI dan pemodelan statistik.
- Big Data Analitik Diagnostik → menemukan jawaban atas alasan lebih mendalam mengenai apa yang sedang terjadi menggunakan drill-down.
- Big Data Analitik Deskriptif → mendapatkan informasi terbaru tentang tren saat ini dan kinerja operasional perusahaan dengan memahami data mentah yang dikumpulkan dan berfokus pada area yang lebih kritis sehingga memahami situasi bisnis di saat ini, mengantisipasi tren masa depan serta membuat keputusan berdasarkan data.

Model Pemrosesan Analitik

Dalam Big Data Analitik, tingkat kesulitan pengolahan datanya semakin besar karena data yang diproses diperoleh dari berbagai sumber dengan bentuk

dan jenis yang berbeda-beda, ukuran dan kecepatan yang besar sehingga dalam prosesnya banyak menggunakan teknik dan algoritma yang lebih advance seperti predictive model dan machine learning untuk melihat trend, pola, korelasi, dan insight lainnya dari data. Selain itu, data analitik seringkali melibatkan teknik dan algoritma pengolahan data yang cukup kompleks seperti data mining maupun perhitungan statistik.

Model pemrosesan data analitik salah satunya dengan menggunakan model prediktif yang berguna untuk memprediksi peristiwa di masa depan dengan menganalisis pola yang berbeda menggunakan data historis yang dilatih, dianalisis, dan dipilih model statistik. Prosesnya melibatkan pelatihan model dimana beberapa model digunakan pada kumpulan data yang sama dan akhirnya didapatkan model yang paling cocok berdasarkan pemahaman data bisnis.

Implementasi Analitik Big Data

- Bidang Transportasi → dengan big data kita dapat memperkirakan potensi kemacetan lalu lintas secara real time, menunjukkan lokasi suatu tempat dalam maps.
- Bidang Kesehatan → membantu memprediksi wabah dan memutuskan tindakan pencegahan untuk meminimalisir efek dari wabah tersebut.
- Bidang Media Sosial → penerapan pada aplikasi Spotify dimana aplikasi akan mengumpulkan data dari user untuk merekomendasikan lagu yang tepat bagi user.

c. Big Data Open Source Tools

- Hadoop → sistem untuk memproses volume data yang sangat besar
- Apache Hbase → sistem yang mampu memproses data dalam skala besar secara interaktif
- Pig → menganalisis data set yang besar dengan menghasilkan kode MapReduce → perhitungan di Hadoop dengan kemampuan memproses data dalam jumlah besar
- Hive → melakukan ekstrak, transformasi, dan load data serta melakukan query terhadap file yang tersimpan di Hadoop Files (HDFS)
- Zookeeper → sebagai sistem distribusi dalam skala besar
- Apache Spark → memproses data dalam skala besar secara in-memory yang dilengkapi API serta mengeksekusi pekerjaan yang membutuhkan perulangan akses secara cepat terhadap data yang diproses.

Referensi:

Naning Nur Wijayanti. (2021, September 23). Apa Itu Big Data? Konsep, Karakteristik, dan Manfaatnya Bagi Bisnis Anda. Retrieved February 18, 2023, from Niagahoster Blog website: https://www.niagahoster.co.id/blog/big-data-adalah/#Apa_Itu_Big_Data

Big data analytics: pengertian, jenis, dan contoh penggunaannya. (2021, September 14).

Retrieved February 18, 2023, from Ekrut.com website:

<https://www.ekrut.com/media/big-data-analytics-adalah>