

Tugas ke-4

Rizka Iqba

19.01.013.116

Tugas Rangkuman Sains Data kelas C

* Sains Data merupakan disiplin ilmu yang secara khusus mempelajari dan menganalisis data. Sains data dapat diterapkan ke berbagai bidang seperti kesehatan, pemerintahan, bisnis, keuangan, pemasaran, dan lain-lain.

* Ada 5 hal yang menganggotakan materi dalam sains data ini yaitu:

- Culture
- Method
- Use Cases
- Research
- Mining

* Data Culture yaitu data yang berada di lingkungan sekitar kita yang jumlahnya sangat besar seperti:

- Astronomi
- Bisnis
- Kedokteran
- Ekonomi
- Olahraga
- Luaca
- Financial
- dll

* Data Mining adalah data yang dikumpulkan lalu diubah menjadi informasi dalam jumlah besar. Dalam proses ini seringkali memanfaatkan metode matematika, statistik, komputer & pemanfaatan teknologi AI (Artificial Intelligence)

* Mining

Data Mining ialah disiplin ilmu yang mempelajari metode untuk menemukan pola, pengetahuan, formula, aturan, ataupun insight dari suatu data yang volumenya besar dan jenisnya beraneka ragam. Konsep *data mining* dimulai dari kumpulan data, lalu *data mining methods*, dan kemudian *knowledge* (pengetahuan)

* Terminologi dan Nama lain dari Data Mining

Ekstraksi dari Data ke Pengetahuan

1. Data: fakta yang terakumulasi tidak membawa arti
2. Informasi: Rekap, rangkuman, penjelasan, visualisasi, dan statistik dari data
3. Pengetahuan: pola, rumus, aturan, atau model yang muncul dari data

* Nama lain Data Mining yaitu sbb:

- Data Science
- Knowledge Discovery in Database (KDD)
- Big Data
- Business Intelligence
- Knowledge Extraction
- Predictive Analytics
- Pattern Analysis
- Information Harvesting

Contoh data di kampus:

- Puluhan ribu data mahasiswa di kampus yang diambil dari sistem informasi akademik
- Seperti apa pengetahuan itu? Rumus, Pola, aturan

* Evolusi Sciences

- Sebelum 1600 : Empirical Science
disebut sains kalau bentuknya kasar mata
 - 1600 - 1950 : Theoretical science
disebut sains kalau bisa dibuktikan secara matematis dan eksperimen
 - 1950 s - 1990 : Computational science
 - Seluruh disiplin ilmu bergerak ke komputasi
 - Lahirnya banyak model komputasi
 - 1990 - sekarang : Data science
 - Kultur manusia menghasilkan data besar
 - Kumpulan komputer untuk mengolah data besar
 - Datangnya data mining sebagai arus utama sains
- * Tugas dan tujuan adanya

Data mining yaitu :

- Mengolah data / dijadikan pola pengetahuan (model / rule / formula / tree)
- Pola pengetahuan dimasukkan ke dalam sistem Software)
- Sistem menjadi cerdas dan bermanfaat signifikan dalam meningkatkan ~~value~~ value perusahaan dan benefitnya

* 4 Dimensi Masalah Big Data yaitu :

- Volume → data preprocessing (Data cleaning, reduction, transformation, dan data integration)
- Variety → Structured vs unstructured data (Text mining) vs (Data mining terstruktur)
- Velocity
- Veracity

* Peran data mining yaitu : estimasi, forecasting, klasifikasi, klastering, dan asosiasi

* untuk Output / pola / Model / Knowledge sendiri yaitu :

- Formula / function (waktu tunggu)
- Decision Tree (pohon keputusan)
- Korelasi dan Asosiasi
- Rule / Aturan
- Cluster

* Penerapan Data science di Organisasi agar berhasil

- Meningkatkan pemahaman dan kompetensi SDM di bidang data mining, paham relevansi
- Kumpulkan dan analisis berbagai database organisasi yang banyak dan memungkinkan etolah menjadi pengetahuan yang bermanfaat

- Gunakan tools dan teknologi yang sederhana dan mudah yang mendukung visi misi organisasi.
- Meningkatkan kompetensi SDM dan pengalaman pengolahan data organisasi agar lebih akurat sesuai kebutuhan.
- * Beberapa platform yang banyak digunakan dalam data mining yaitu : KNIME , Rapidminer , dll
- * Private and Sektor Komersial, yakni :
 - Marketing , Finance , Banking and Insurance , Keamanan , Manufacturing , Web dan Internet , Software Engineering , dan Telecommunication
- * Adapun public and Government sector yaitu :
 - Finance , taxation , Medicine and Health Care , Education , Insurance , Security , Transportation , Law , and politic
- * Perbedaan Gaya Penelitian Data Science

Computing Researcher	Non Computing Researcher
<ul style="list-style-type: none"> • Perbaiki algoritma dan metode • Pengukuran performa dari Algoritma 	<ul style="list-style-type: none"> • Analisis pola yang dihasilkan algoritma • multi aspect dan perspective (Finance, Energi, Hukum, dll) • Pengukuran ke manfaat pola untuk kehidupan