



Type Individu

Resume Minggu [Ke - 9]

[Dodi Devrian Andrianto]

[119140023]

[dodi.119140023@student.itera.ac.id]

[26 Oktober 2022]

Ketentuan

1. Adanya referensi tambahan untuk tugas kuliah tamu
2. Tidak Boleh sama antar mahasiswa
3. Minimal 2 Halaman dengan ketentuan A4, Margin 3x3x3x3 dan spasi 1,5

Topik

Data Science

KONSEP DATA SCIENCE

- Pada dasarnya, DS menggali dan mengekstrak "knowledge" dan "insight" dari data.
- Di era digital ini, hampir semua kegiatan manusia, interaksi manusia dengan lingkungannya, dan lingkungannya itu sendiri, bisa terekam dalam bentuk digital, atau bisa di digitalisasikan, sehingga menjadi data.
- Sehingga dapat dikatakan "everything is data".
- Secara inherence, ternyata berbagai data tersebut ada strukturnya (inherence structure, atau hidden structure).
- Menjadi tantangan tersendiri bagi data scientist untuk mengungkapkan inherence structure, atau hidden structure tersebut, menjadi "knowledge" and "insight".

TECHNIQUES

- classification: predict class from observations
- clustering: group observations into "meaningful" groups
- regression (prediction): predict value from observations



**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN,
RISET, DAN TEKNOLOGI
INSTITUT TEKNOLOGI SUMATERA
JURUSAN TEKNOLOGI PRODUKSI DAN INDUSTRI**

Jalan Terusan Ryacudu Way Hui, Kecamatan Jati Agung, Lampung Selatan 35365
Telepon: (0721) 8030188
Email: jtpi@itera.ac.id, Website : <http://itera.ac.id>

POPULAR FRAMEWORKS/TOOLS

- Weka
- Carrot2
- Gate
- OpenNLP
- LingPipe
- Stanford NLP
- Mallet - Topic Modelling
- Gensim – Topic Modelling (Python)
- Apache Mahout
- MLib – Apache Spark
- scikit-learn – Python
- LIBSVM : Support Vector Machines

KEMAMPUAN PEMROGRAMAN

Salah satu kemampuan utama yang dibutuhkan sebagai data scientist adalah dapat melakukan pemrograman dan menggunakan tools yang membantu dalam menganalisis data. Kamu juga harus memiliki pengetahuan dan pengalaman coding dengan beberapa bahasa seperti R, Python, dan lainnya. Adapun pemrograman dengan penggunaan bahasa seperti python ini berfungsi untuk memudahkan para data scientist untuk mengatur atau mengorganisir kumpulan data yang tidak terstruktur. Jika kamu ingin mempelajari dasar-dasar pemrograman tersebut, kamu bisa langsung mengikuti kelasnya dengan mengklik banner di bawah ini.

KEMAMPUAN DATABASE, QUERY (SQL), DAN PENGOLAHAN DATA

Seorang data scientist membutuhkan SQL (Structured Query Language) untuk menangani data terstruktur yang ada pada database sehingga menjadi lebih mudah. Adapun hal ini juga termasuk pemahaman mengenai SQL commands seperti,

- Data Query Language
- Data Manipulation Language



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN,
RISET, DAN TEKNOLOGI

INSTITUT TEKNOLOGI SUMATERA

JURUSAN TEKNOLOGI PRODUKSI DAN INDUSTRI

Jalan Terusan Ryacudu Way Hui, Kecamatan Jati Agung, Lampung Selatan 35365

Telepon: (0721) 8030188

Email: jtpi@itera.ac.id, Website : <http://itera.ac.id>

- Data Definition Language
- Data Control Language

[*Dodi Devrian Andrianto*]