



Text Mining

Introduction

pretest1 : <https://tinyurl.com/pretest1tm>



follow :  tin_uinam



| aboutme



nama : Muhammad Nur Akbar | MNA
pendidikan : S1 Institut Teknologi Telkom
S2 Institut Teknologi Bandung
profesi : Dosen Mata Kuliah Data Mining
ex Data Analyst in Traveloka for International Growth
WA : 0852 4204 4250
email : muhammad.akbar@uin-alauddin.ac.id



| Outline

1. Kontrak perkuliahan
2. Deskripsi singkat
3. Capaian prodi dan mata kuliah
4. Silabus perkuliahan Text Mining
5. Introduction to Text Mining



| SEMUA MAHASISWA NILAI **A**, So...

Tips Agar Nilai **TIDAK** A

- Tidak usah masuk kelas
- Tidak perlu mengerjakan tugas dengan maksimal
- Tidak perlu belajar untuk UTS
- Tidak perlu mengerjakan Tugas Besar
- Silahkan titip absen
- Silahkan nyontek dan curang



| #1 Kontrak Perkuliahan

- 3 SKS :
 - 3 x 50 menit tatap muka / pekan
 - 3 x 60 menit tugas mandiri / pekan
- Kompetensi akan diuraikan beberapa komponen :
 - Kehadiran : 5%
 - UTS : 20%
 - PreTest, PostTest, Quiz : 30%
 - Tugas Mandiri : 15%
 - UAS (Tugas Besar) : 30%

A = 86-100

B = 76-85

C = 66-75

D = 56-65

E = 0-55



| #1 Kontrak Perkuliahan (cont)

- Kehadiran minimal 80% (13 dari 16 kehadiran)
- Tidak ada quiz, tugas, PreTest, PostTest susulan
- Tidak ada tugas tambahan untuk perbaikan nilai
- Plagiat atau titip presensi **auto E**
- UTS atau UAS susulan hanya untuk **kasus khusus**
- Tidak diperbolehkan mengenakan kaos, sandal, serta celana bagi mahasiswi
- Toleransi keterlambatan **20 menit**



| #2 Deskripsi Singkat

- Text mining merupakan penerapan konsep dan teknik data mining untuk mencari pola dalam teks
 - Proses penganalisisan teks guna menyarikan informasi yang bermanfaat untuk tujuan tertentu.
- Proses data mining untuk data dokumen atau teks memerlukan lebih banyak tahapan, mengingat data teks memiliki karakteristik yang lebih kompleks daripada data biasa.

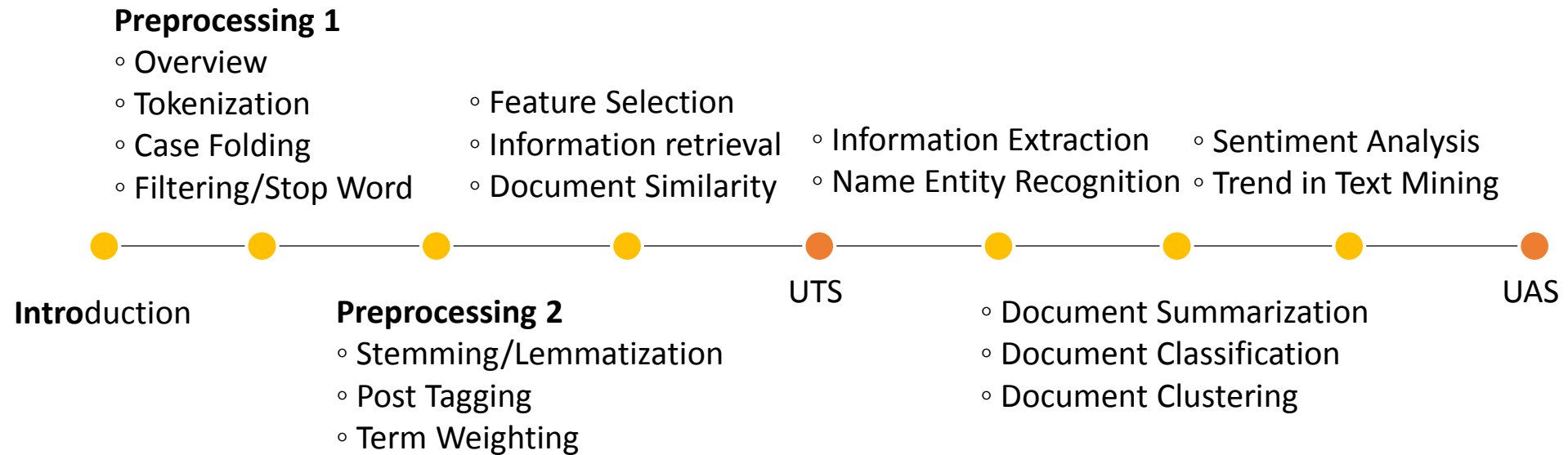


| #3 Capaian Prodi dan Mata Kuliah

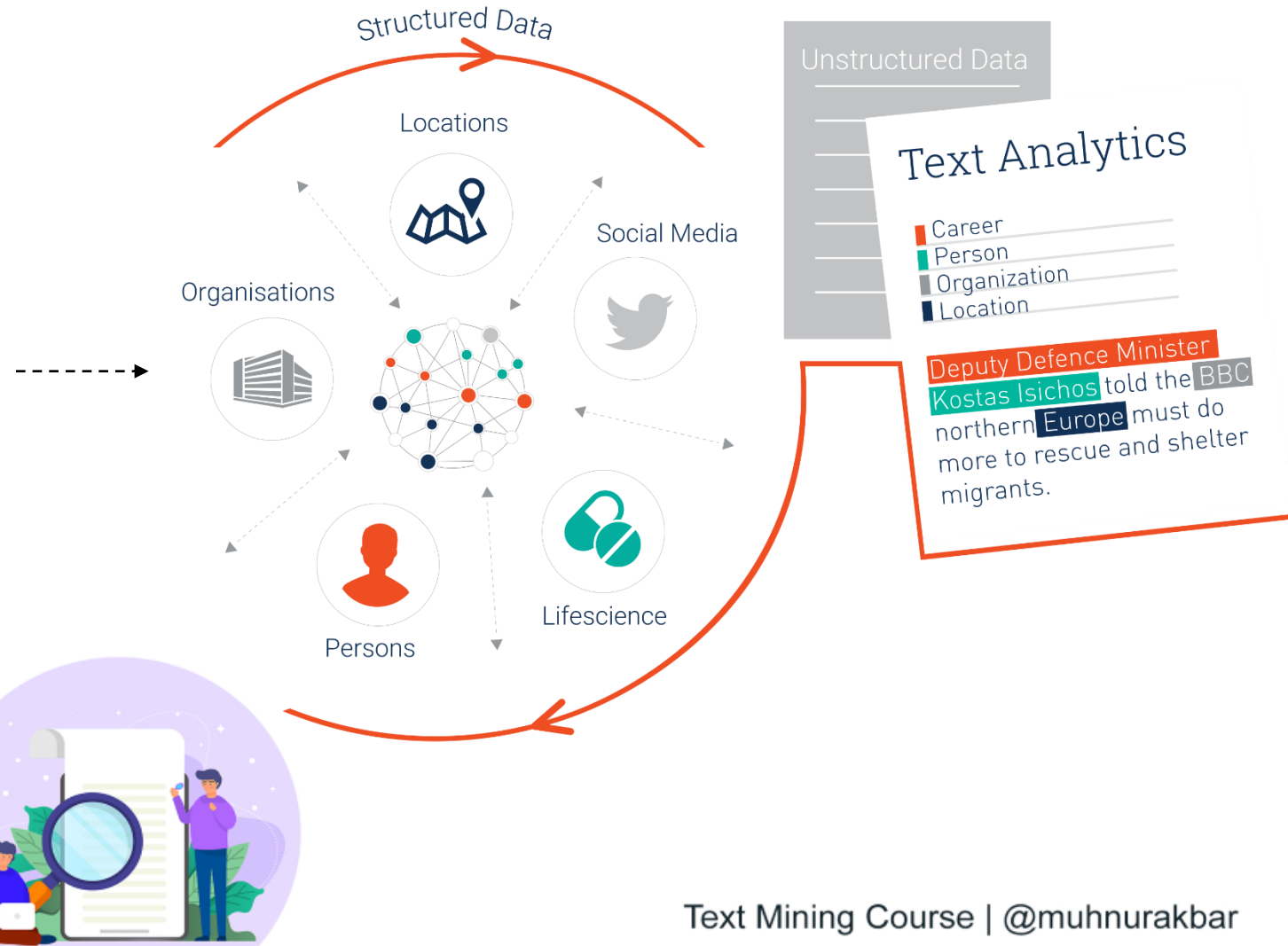
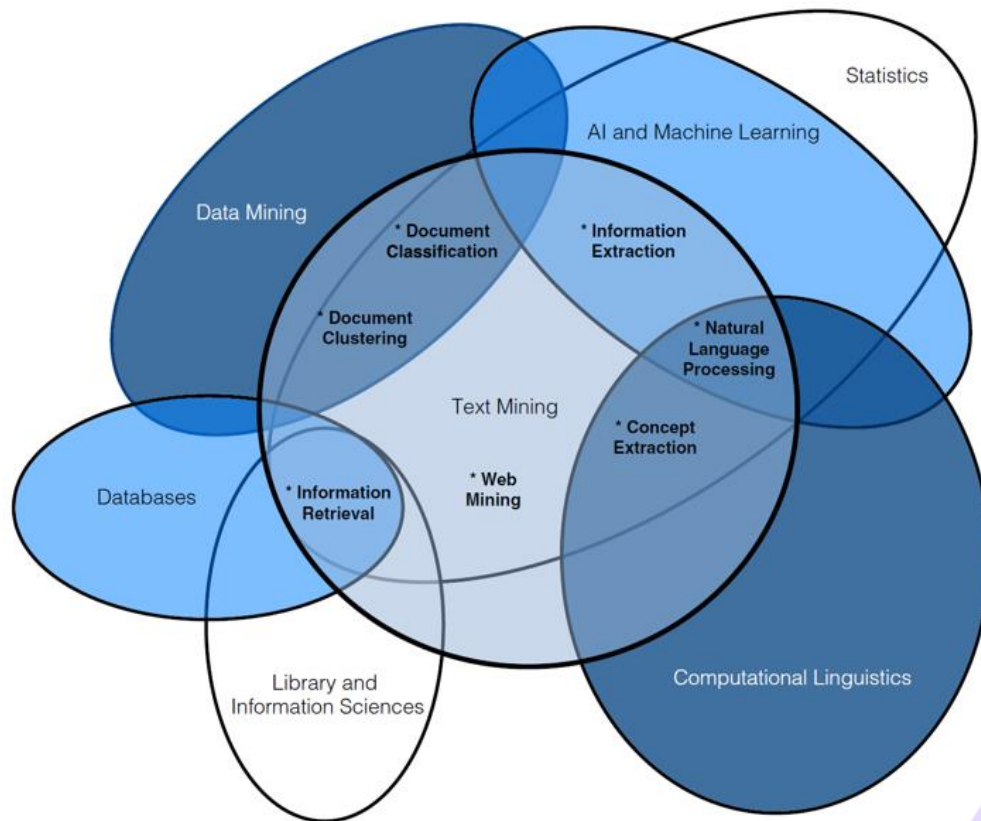
- Capaian Prodi:
 - Mampu menggunakan metode knowledge discovery yang tepat untuk menganalisis data text/document
 - Menguasai konsep teoritis Text Mining dan mengaplikasikan dalam kasus riil
- Capaian Mata Kuliah:
 - Mampu menganalisa kebutuhan data sesuai kebutuhan bisnis
 - Mempersiapkan dan mengolah data untuk keperluan analisa
 - Menganalisis teori-teori dasar dan algoritma Text Mining meliputi Information Retrieval, Information Extraction, Summarization, Document Clustering, Document Classification dalam penerapannya pada kasus riil
 - Mampu menyelesaikan tugas matakuliah tepat waktu secara mandiri atau berkelompok



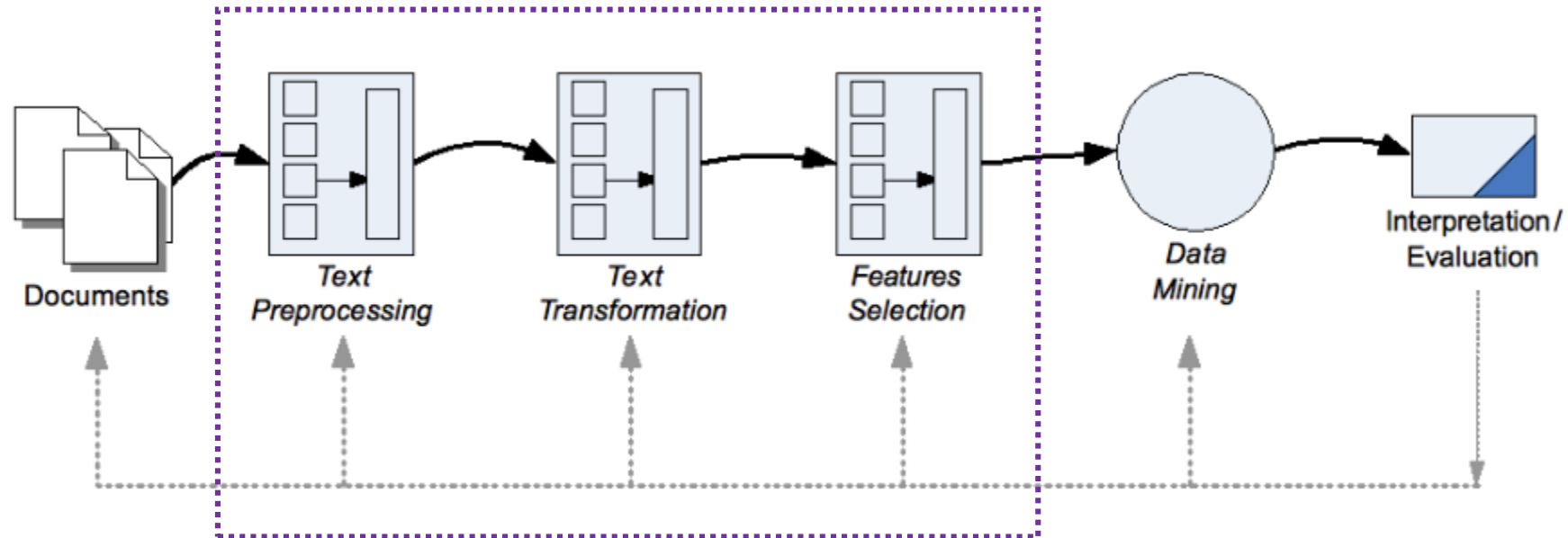
| #4 Silabus Perkuliahan



| #5 Introduction to Data Mining (cont)

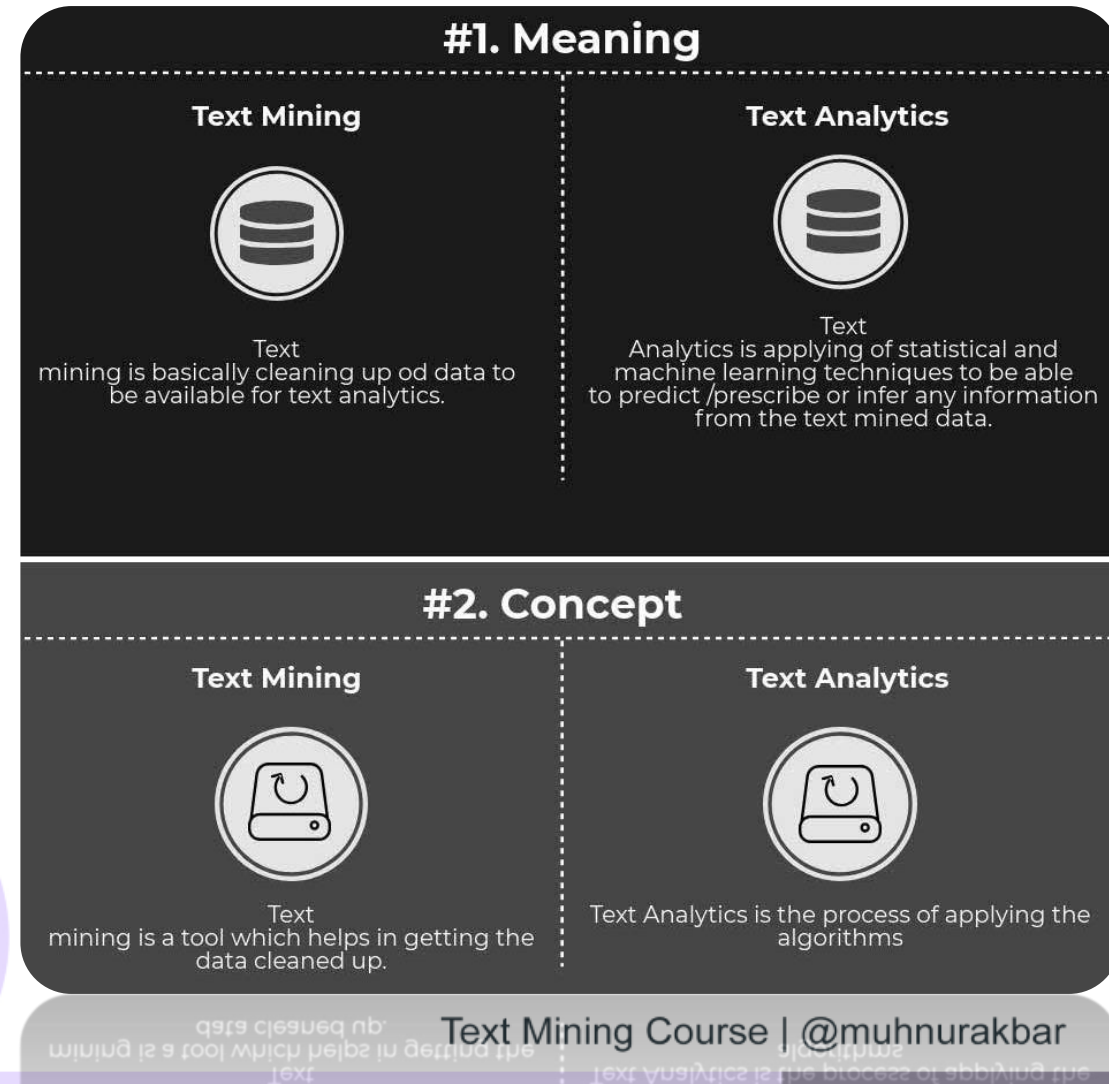


| #5 Introduction to Data Mining (cont)



| #5 Introduction to Data Mining (cont)

- Apa yang dimaksud *text* di sini?
 - Teks bahasa manusia
 - Teks dapat berupa (granularity): kata, term, frase, kalimat, dokumen, ..
- Apa yang dimaksud dengan *mining* di sini?
- Text mining \approx Text analytics



| #5 Introduction to Data Mining (cont)

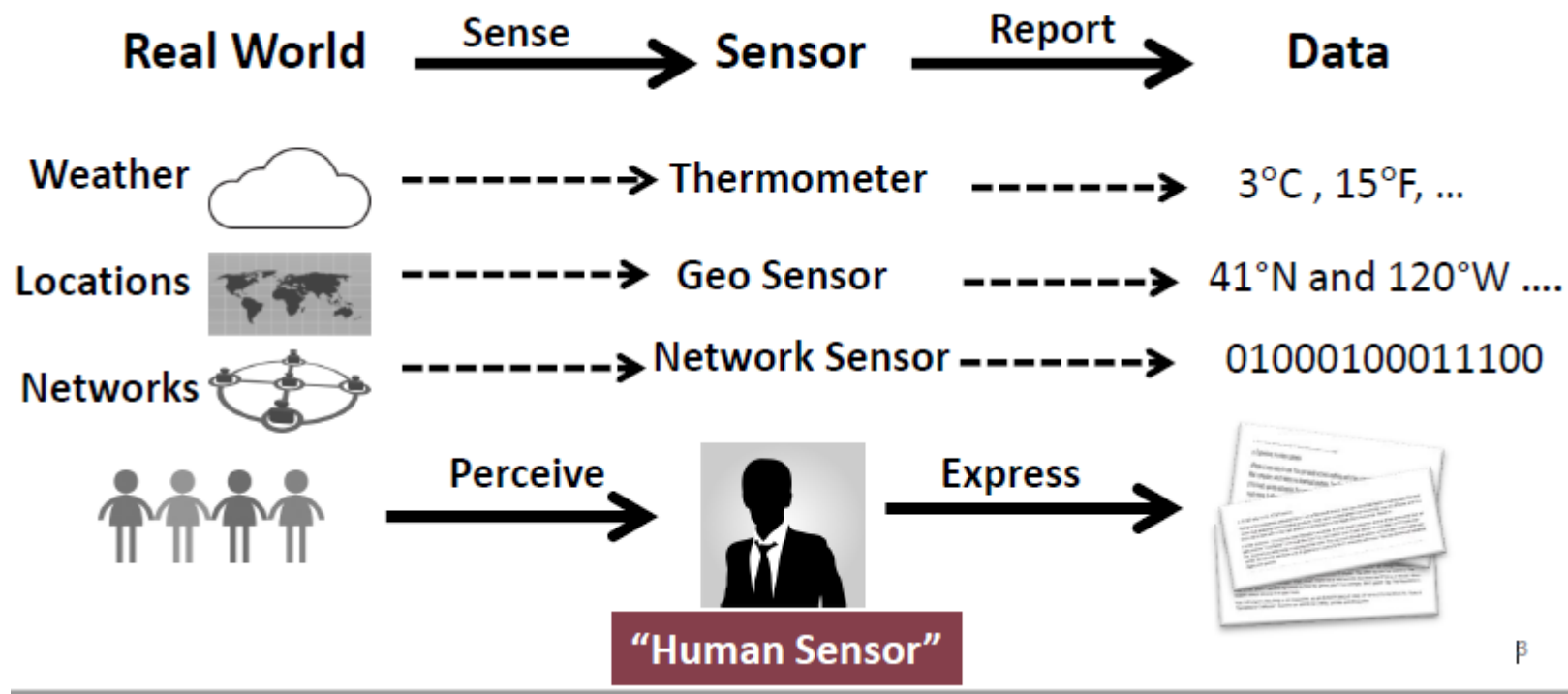
Karakteristik Teks

- Tdk terstruktur/semi terstruktur.
- Dimensi tinggi.
- Penekanan pada makna.



| #5 Introduction to Data Mining (cont)

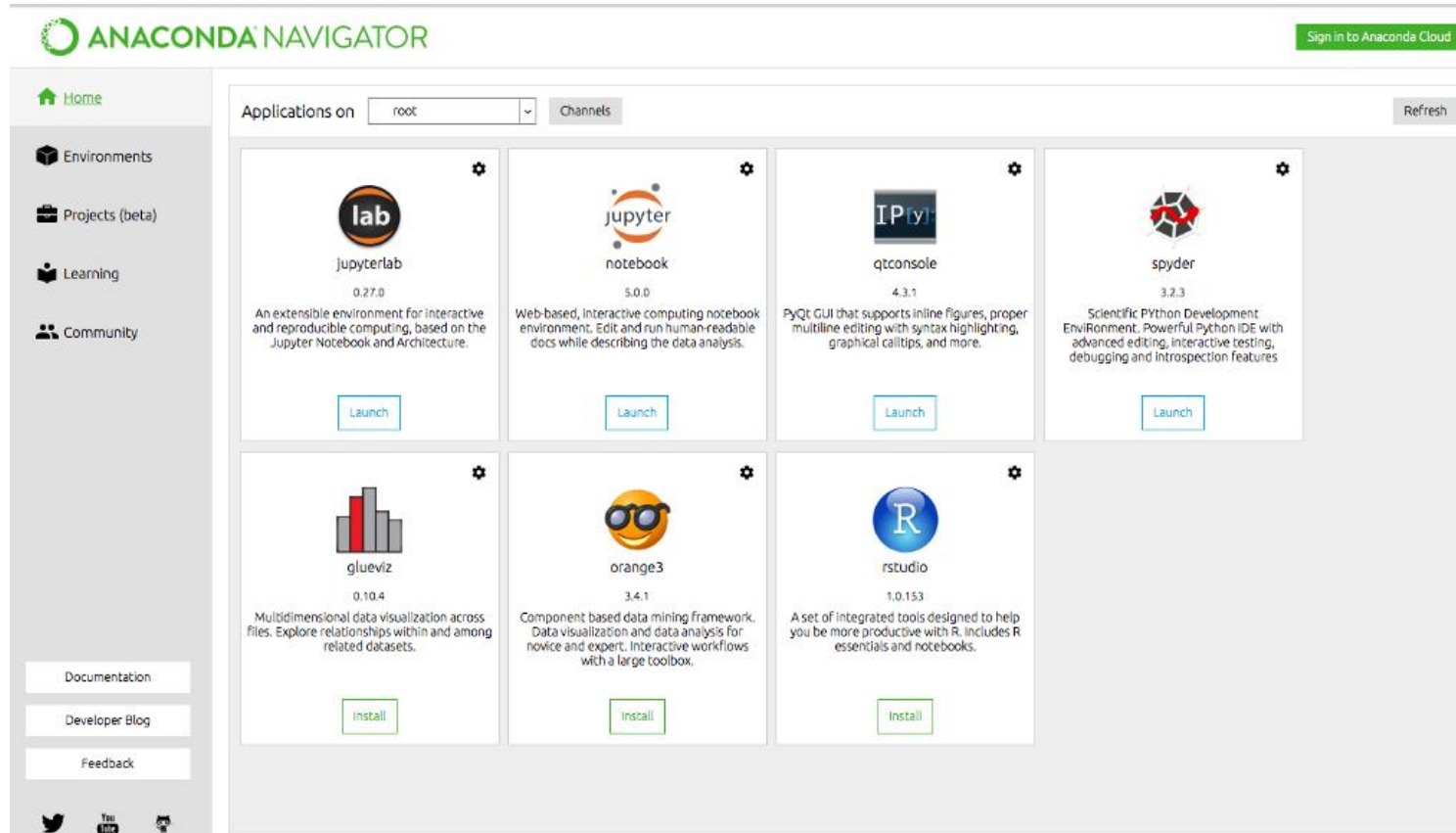
- Manusia sebagai “sensor” yg subyektif



Dari: Zhai, "Text Mining and Analytics course", Illinois Univ

#TASK1 – Google Classroom

- Install **Orange** + screenshot
<https://orangedatamining.com/download/>
- Run “Hello world!” [> Py 3.0] di **Jupyter Notebook** + screenshot



“You can dig out facts, but **if you don't go deep enough**, the facts won't tell you much.”

(Jeff Catlin, CEO Lexalytics Inc)



Thank you

