



CHULA **ENGINEERING**
Foundation toward Innovation

COMPUTER

Course Overview

2110531 Data Science and Data Engineering Tools

Assoc. Prof. Peerapon Vateekul, Ph.D.
Department of Computer Engineering,
Faculty of Engineering, Chulalongkorn University
Peerapon.v@chula.ac.th
www.cp.eng.chula.ac.th/~peerapon/

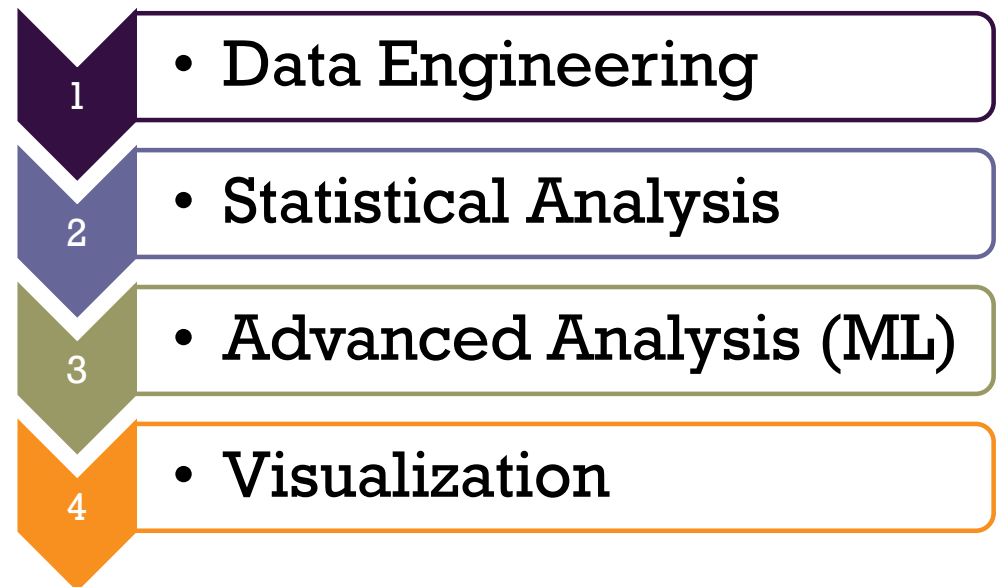
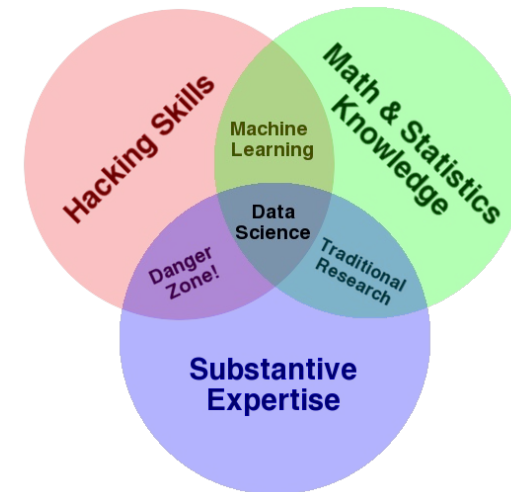
+ What is this class about? (recap)

Dr.Virote

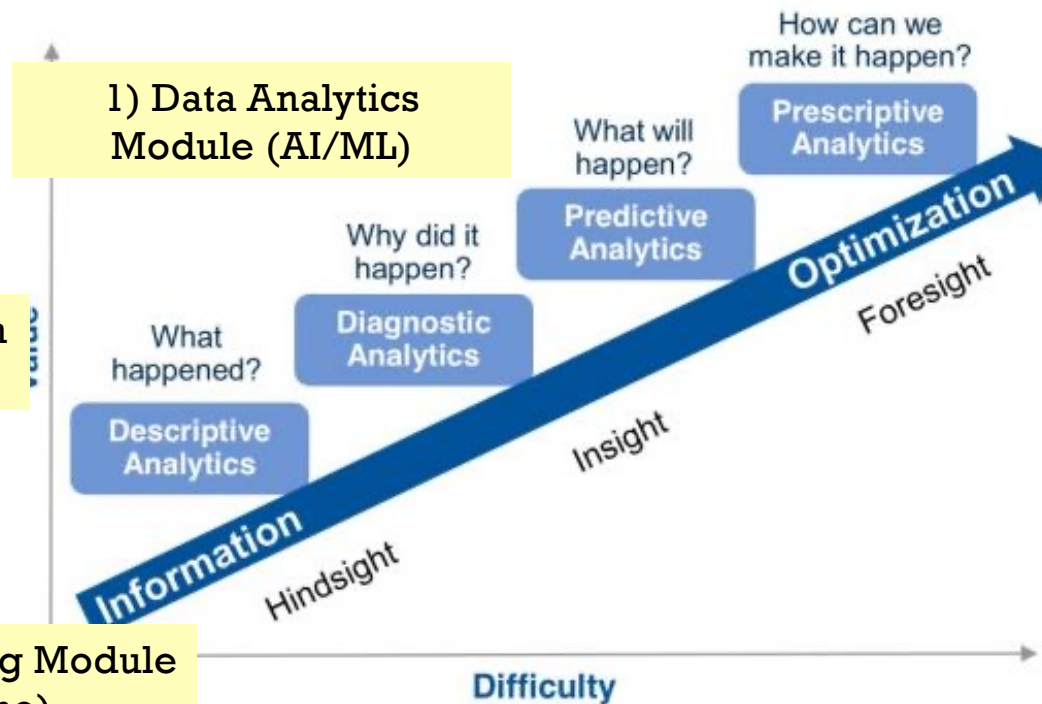
1. Transform data into **valuable insights**
2. Transform data into **data products**
3. Transform data into **interesting stories**

Aj.Natawut

1. Measurement
2. Insights
3. Data Products



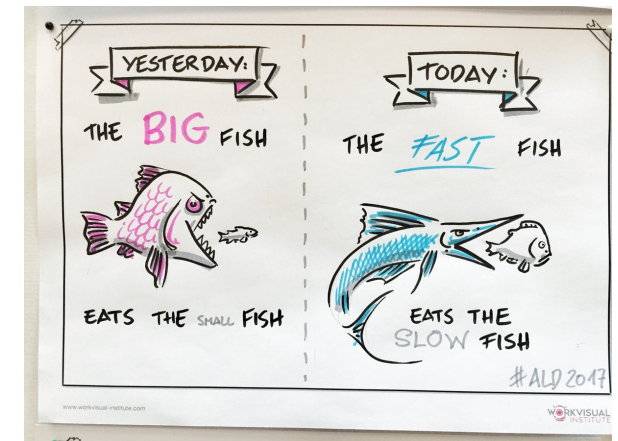
What is this class about? 3 Modules



Gartner's Top Strategic Technology Trends for 2023

Optimize	Scale	Pioneer
Digital Immune System	Industry Cloud Platforms	Superapps
Applied Observability	Platform Engineering	Adaptive AI
AI TRISM	Wireless-Value Realization	Metaverse
Sustainable Technology		

Gartner





Class tentative schedule

- Section1 on Tue 1PM-4PM, Building ENG3 Room 417
- Section5 on Sat 1PM-4PM, Building ENG4 Room 18-16

#	Tue 1PM-4PM	Sat 1PM-4PM	Contents	Instructor	Module
1	08-Aug-23	12-Aug-23 Online	Introduction, Pandas, Data Prep (12-Aug is Mother Day)	Aj.Peerapon	Data Science
2	15-Aug-23	19-Aug-23	Traditional ML (1)	Aj.Peerapon	Data Science
3	22-Aug-23	26-Aug-23	Traditional ML (2)	Aj.Peerapon	Data Science
4	29-Aug-23	02-Sep-23	Deep Learning (1); CNN, RNN (LSTM, GRU)	Aj.Peerapon	Data Science
5	05-Sep-23	09-Sep-23	Deep Learning (2); Transformer	Aj.Peerapon	Data Science
6	12-Sep-23	16-Sep-23	Advanced topics (Generative AI) + Model monitoring (Mlflows)	Aj.Peerapon	Data Science
7	19-Sep-23	23-Sep-23	Big data architecture + data storage	Aj.Natawut	Big Data Eng.
	26-Sep-23	30-Sep-23	Midterm Exam Week (25 - 29 Sep 2023)		
8	03-Oct-23	07-Oct-23 Online	Web scraping (7-Oct is Graduation Day)	Aj.Natawut	Big Data Eng.
9	10-Oct-23	14-Oct-23	Data ingestion	Aj.Natawut	Big Data Eng.
10	17-Oct-23	21-Oct-23	Big data processing (Spark)	Aj.Natawut	Big Data Eng.
11	24-Oct-23	28-Oct-23	MLOps: Orchestration (Airflow) and serving (FastAPI, Seldon Core)	Aj.Natawut	Big Data Eng.
12	31-Oct-23	04-Nov-23	Guest speakers (AWS)	Aj.Peerapon	Data Science
13	07-Nov-23	11-Nov-23	Data visualization	Aj.Veera	Data Viz
14	14-Nov-23	18-Nov-23	Python visualization	Aj.Veera	Data Viz
15	21-Nov-23	25-Nov-23	Graph analysis & spatial analysis	Aj.Veera	Data Viz
	28-Nov-23	02-Dec-23	Final Exam Week (27 Nov - 12 Dec 2023) *** Final Exam on Sat 2 Dec 2023		

* There will be up to two guest speakers in the class.



Evaluation

■ Module1 Assignment (data analytics)	15%
■ Module2 Assignment (data engineering)	15%
■ Module3 Assignment (data visualization)	10%
■ Midterm Exam (Kaggle)	15%
■ Project	15%
■ Final Exam	30% (Lab Test)



Class communication

- CourseVille: “GenerativeAI”
- Discord: <https://discord.gg/CgVwVtnuh2>
- GitHub: https://github.com/pvateekul/2110531_DSDE_2023s1



Remarks

- **This course assumes that:**
 - It is **just a fundamental course** to perform a data science task properly. Then, you can choose your own specialty in AI domains.
 - It aims to utilize all tools and algorithms **without** theoretical proofs.
 - All students should have a background in **Python** programming.
 - This course (2110531) is designed **only for graduate students**. There is another course (2110446) designed for undergraduate students.
- **There are many courses related to AI/ML/DS?**
 - AI
 - Data Mining
 - Machine Learning
 - Neural Networks
 - ASR
 - NLP
 - Computer Vision
 - Etc.



Any questions? 😊