Course Overview

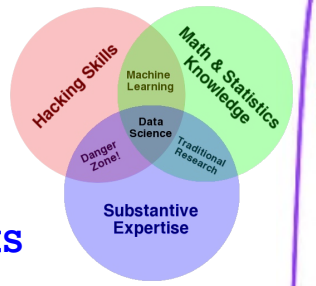
2110446 Data Science and Data Engineering (2023/2)

CHULA ENGINEERING
Foundation toward Innovation
COMPUTER

* Part of this slide is modified from a slide of Prof.Natawut

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)

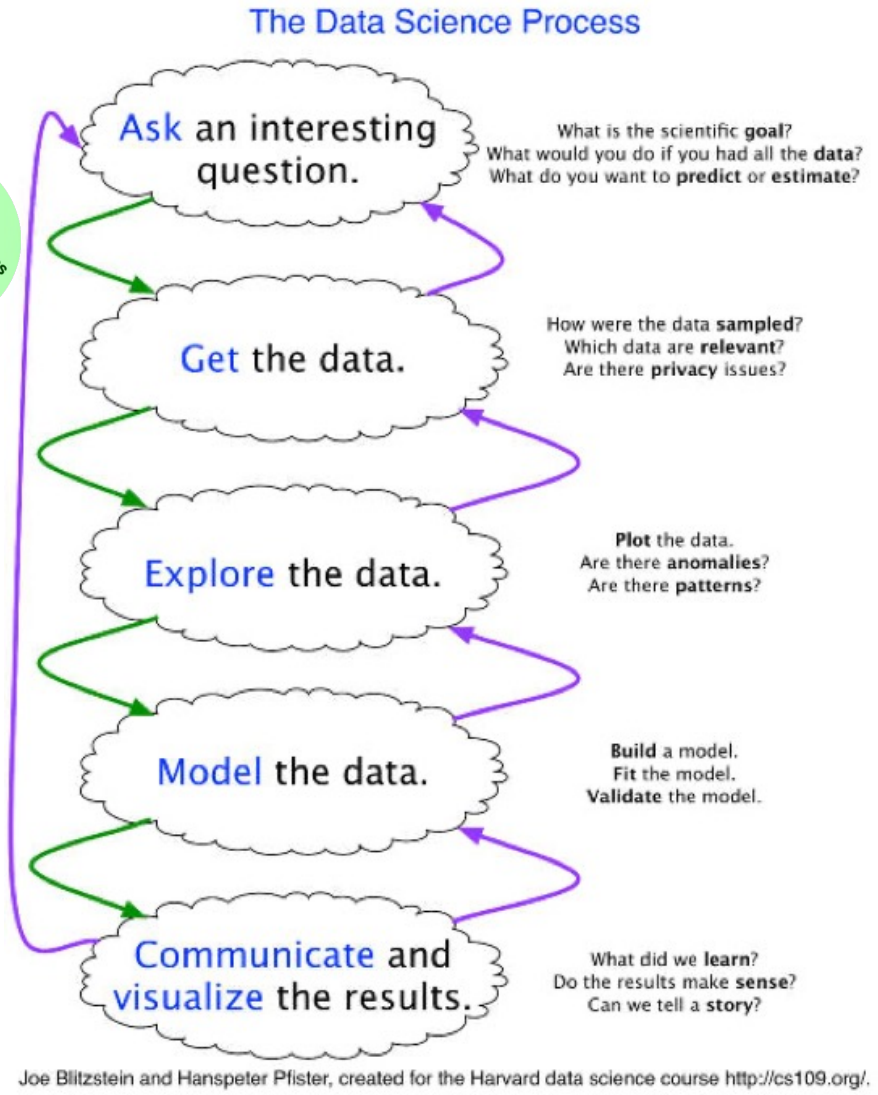


Dr.Virote

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

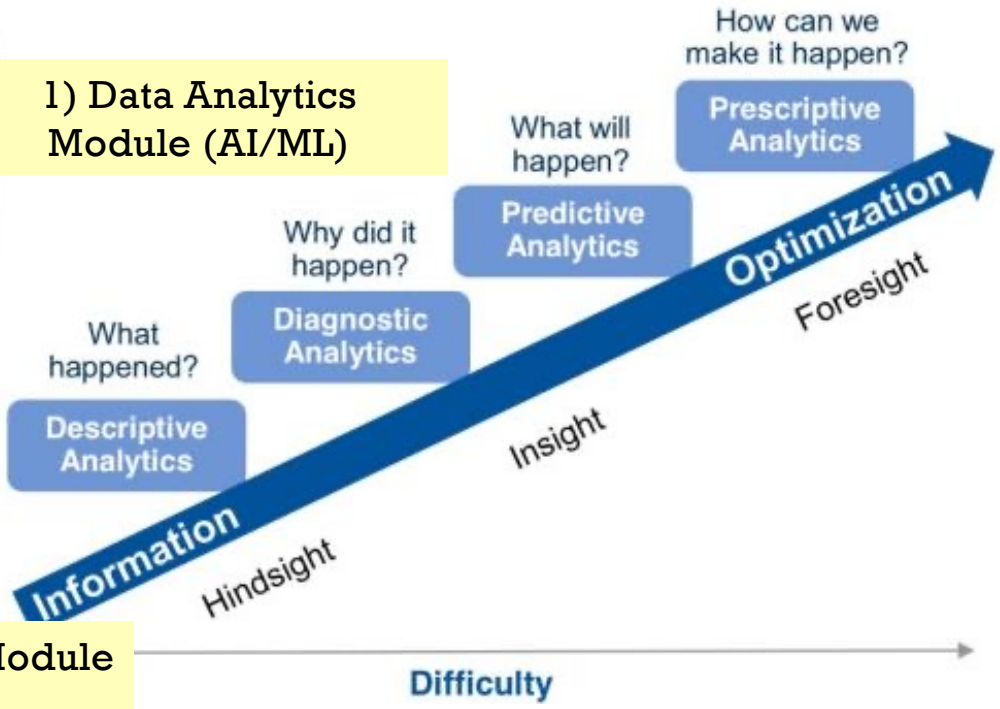
Aj.Natawut

1. Measurement (decision)
2. Insights (knowledge)
3. Data Products (Innovation, Intelligent)

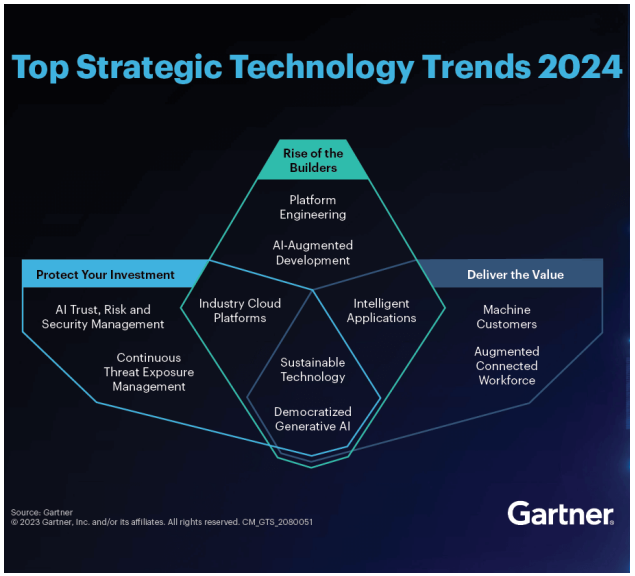


What is this class about? 3 Modules

1) Data Analytics Module (AI/ML)



3) Data Vizualiation Module



2) Data Engineering Module (Data Pipeline)

#	Tue (1PM-4PM)	Contents	Instructor	Module
1	9-Jan-24	Introduction & data preparation	Aj.Peerapon	Data Science
2	16-Jan-24	Traditional ML (1)	Aj.Peerapon	Data Science
3	23-Jan-24	Traditional ML (2)	Aj.Peerapon	Data Science
4	30-Jan-24	DL (1)	Aj.Peerapon	Data Science
5	6-Feb-24	DL (2)	Aj.Peerapon	Data Science
6	13-Feb-24	Advanced topics & Guest speakers * Modern ML techniques (Transformer) * Sharing data scientist experience (in Thailand & Abroad) [2-3 guest speakers]	Aj.Peerapon	Data Science
7	20-Feb-24	Big data architecture + data storage	Aj.Natawut	Big Data Eng.
8	27-Feb-24	Web scraping	Aj.Natawut	Big Data Eng.
	5-Mar-24	Midterm Exam Week (4-8 Mar)		
9	12-Mar-24	Data ingestion	Aj.Natawut	Big Data Eng.
10	19-Mar-24	Big data processing (Spark)	Aj.Natawut	Big Data Eng.
11	26-Mar-24	MLOps: Orchestration (Airflow) and serving (FastAPI, Seldon Core)	Aj.Natawut	Big Data Eng.
12	2-Apr-24	Data visualization	Aj.Veera	Data Viz.
13	9-Apr-24	Python visualization	Aj.Veera	Data Viz.
14	16-Apr-24	Graph analysis & spatial analysis (online)	Aj.Veera	Data Viz.
15	23-Apr-24	Final exam		
	30-Apr-24	Final Exam Week (29 Apr -14 May)		

Evaluation

- Module1 Assignment (data analytics) 15%
- Module2 Assignment (data engineering) 15%
- Module3 Assignment (data visualization) 10%
- Midterm Exam (Kaggle) 15%
- Project 10%
- Final Exam 30% (Lab Test)
- Attendance 5%
  - Students need to attend the class on-site at least 80% (at least 12 weeks) as a mandatory criterion to "pass" this course.

Class communication

- myCourseVille: “GenerativeAI”
- Discord: https://discord.gg/XzDbZcGmT2
- Github: https://github.com/pvateekul/2110446_DSDE_2023s2

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.

ปีการศึกษา 2566/ภาคการศึกษาปลาย
ทวิภาค
2110446 DATA SCI/ENG
วิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูล
DATA SCIENCE AND DATA ENGINEERING
คณะวิศวกรรมศาสตร์ (ภาควิชาวิศวกรรมคอมพิวเตอร์)
3.0 CREDIT HOURS = LECT 3.0 CR
(LECT 3.0 HR)
เงื่อนไขรายวิชา : -
วันสอบกลางภาค : TDF (รอประกาศ) วันสอบปลายภาค : TDF (รอประกาศ)

ตอนเรียน	วิธีสอน	วัน-เวลาเรียน	อาคาร	ห้อง	ผู้สอน	หมายเหตุ	จำนวนนิสิต Regis/Max
21	LECT	TU 13:00-16:00	EN100	201A	NNP,VMS,PVK	APPROVED CP	100/100

- **There are many courses related to AI/ML/DS?**
- AI
- Data Mining
- Machine Learning
- Neural Networks
- ASR
- NLP
- Computer Vision
- Etc.

+

Any questions? ☺