Lessons Learned from Running Interdisciplinary Computational Sciences Program for Undergraduate at Seoul National University on the Future of Scientific Computing Education

Sang-Mook Lee

School of Earth and Environmental Sciences

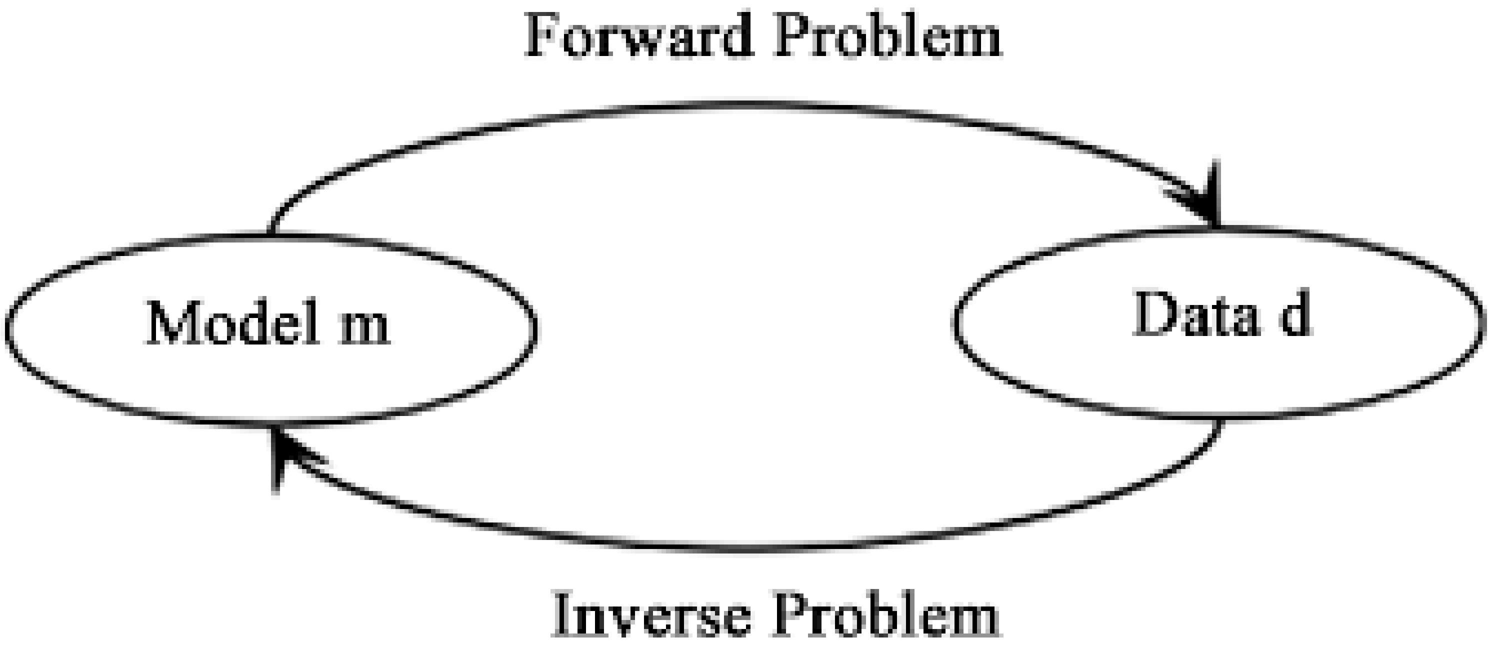
Interdisciplinary Program on Computational Sciences (Undergraduate and Graduate, 계산과학 연합전공 및 협동과정)

Seoul National University

Outline

- A little bit about myself
  - Geophysical inversion (overdetermined, underdetermined system) Singular Value Decomposition
  - Numerical Analysis (Calculus and Linear Algebra)
- AI, Machine Learning, and Deep Learning
- But there are always not enough data (observations)
- Scientific Computing (traditional) versus Enterprise Data Science(emerging)
- How should we approach Computational Sciences Education?

What is the inverse problem (inversion)?



$$A\,\mathbf{x} = \mathbf{b} \qquad\qquad A\,\mathbf{x} = \lambda\,\mathbf{x}$$

Linear(-ized) inverse problems can be formulated in the following way:

$$\boxed{d_i = G_{ij} m_j}$$

(summation convention applies)

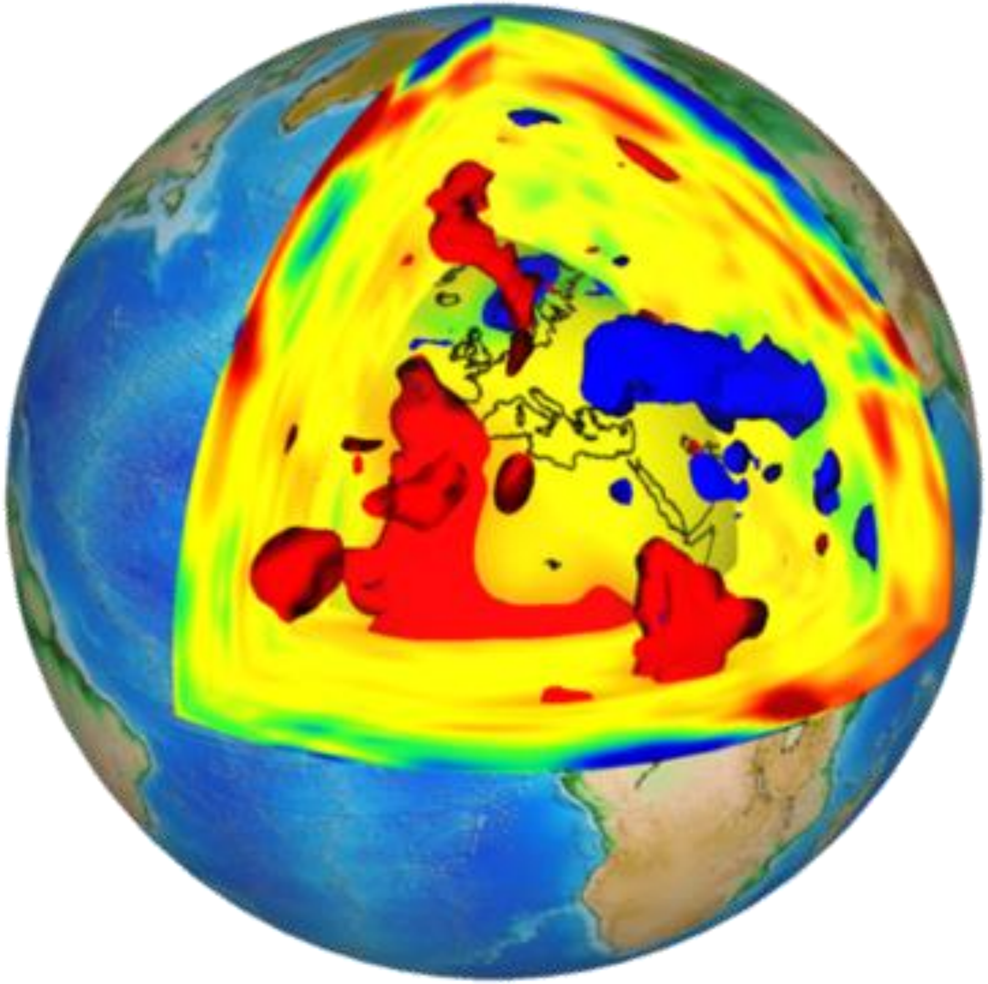
i=1,2,…,N	number of data	
j=1,2,…,M	number of model parameters	
G_{ij}	known	(mxn)

We observe:

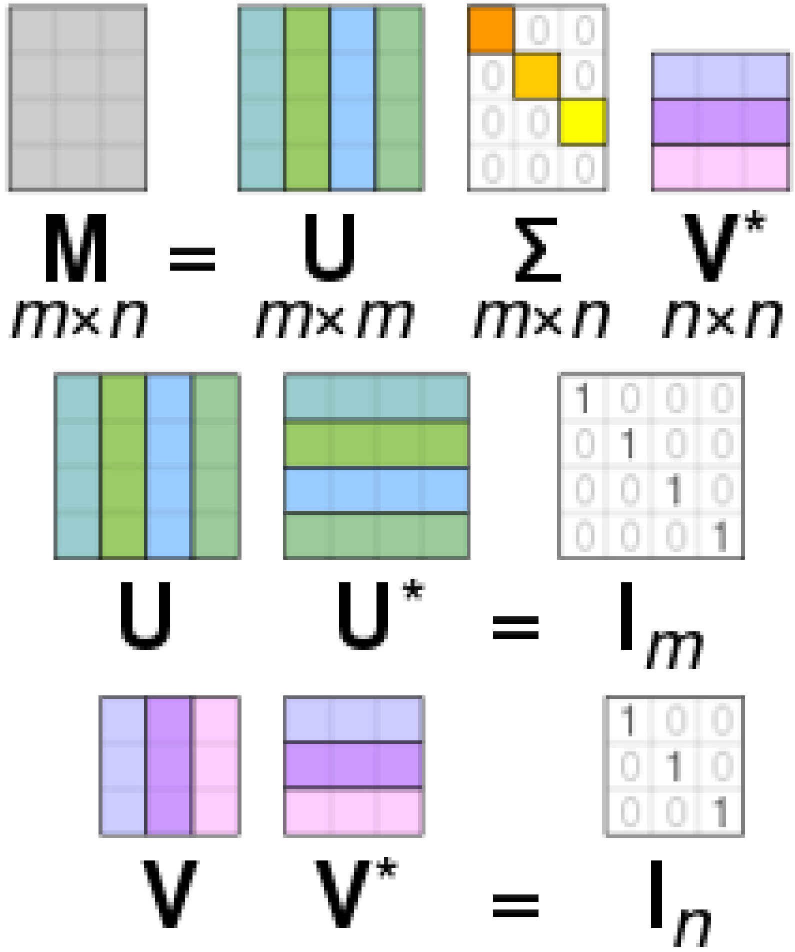
- The inverse problem has a unique solution if N=M and det(G)≠0, i.e. the data are linearly independent
- the problem is overdetermined if N>M
- the problem is underdetermined if M>N

Taught under the name of REGRESSION in ML

Global Seismic Tomography



Singular Value Decomposition



Taught under the name of PCA in ML

Then all of a sudden, *Machine Learning and Deep Learning*

REVIEW

GEOPHYSICS

Machine learning for data-driven discovery in solid Earth geoscience

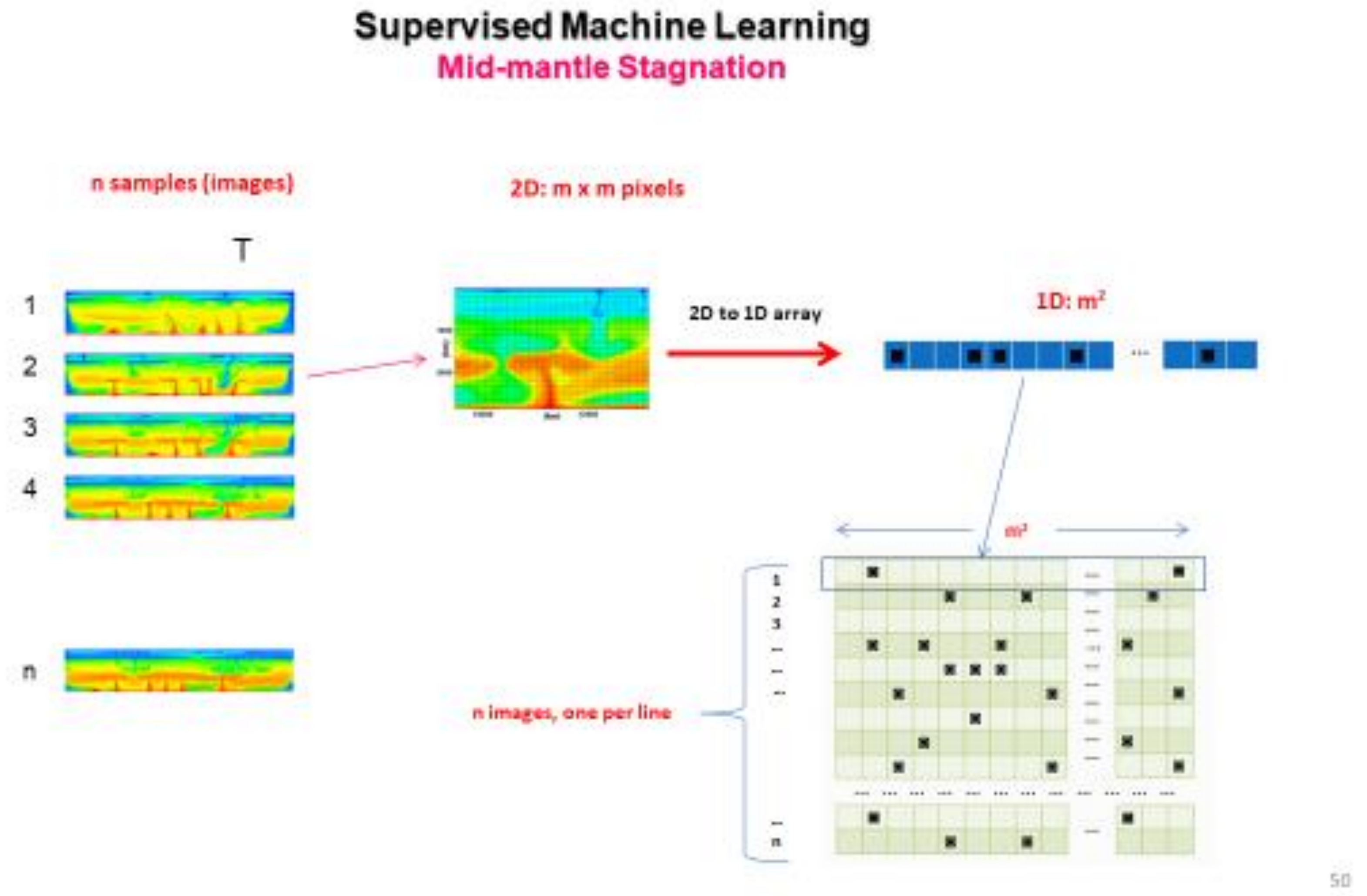
Karianne J. Bergen[1,2], Paul A. Johnson[3], Maarten V. de Hoop[4], Gregory C. Beroza[5]*

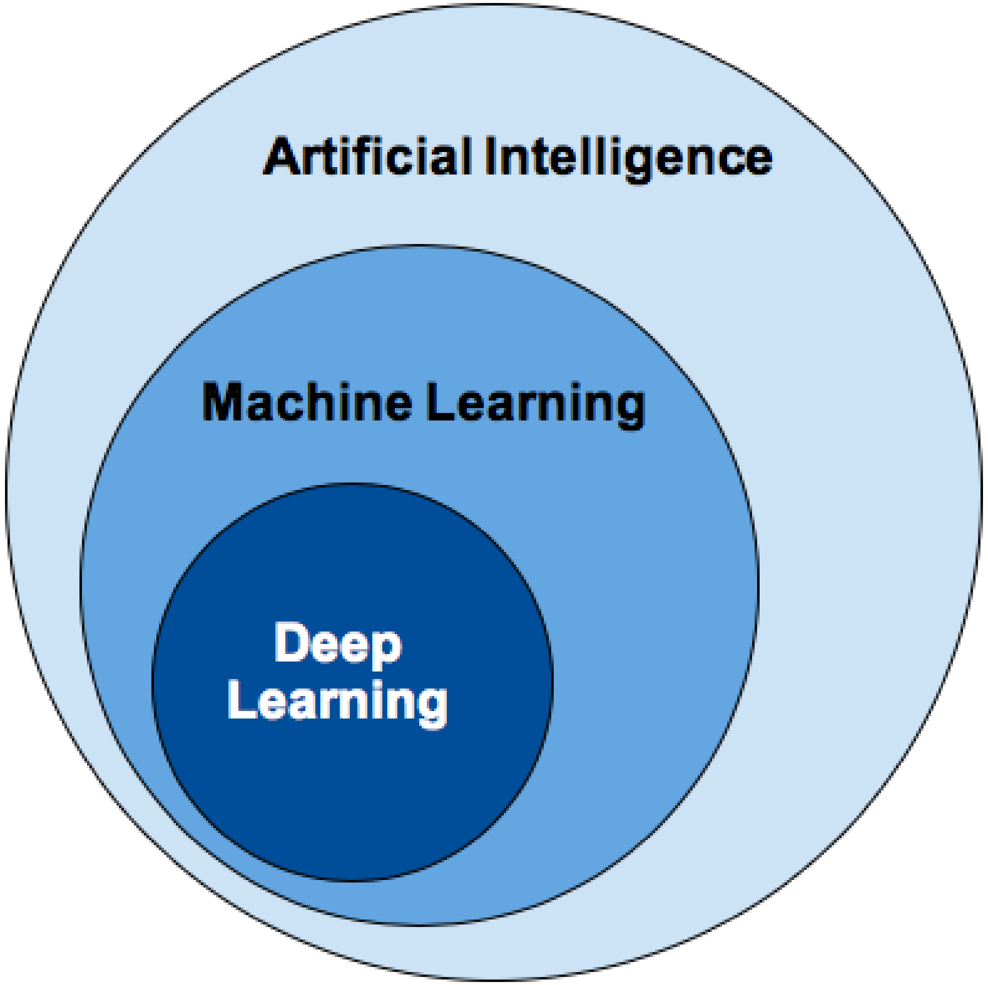
Understanding the behavior of Earth through the diverse fields of the solid Earth geosciences is an increasingly important task. It is made challenging by the complex, interacting, and multiscale processes needed to understand Earth's behavior and by the inaccessibility of nearly all of Earth's subsurface to direct observation. Substantial increases in data availability and in the increasingly realistic character of computer simulations hold promise for accelerating progress, but developing a deeper understanding based on these capabilities is itself challenging. Machine learning will play a key role in this effort. We review the state of the field and make recommendations for how progress might be broadened and accelerated.

Science 2019

An Example of Machine Learning for Geodynamical Problems

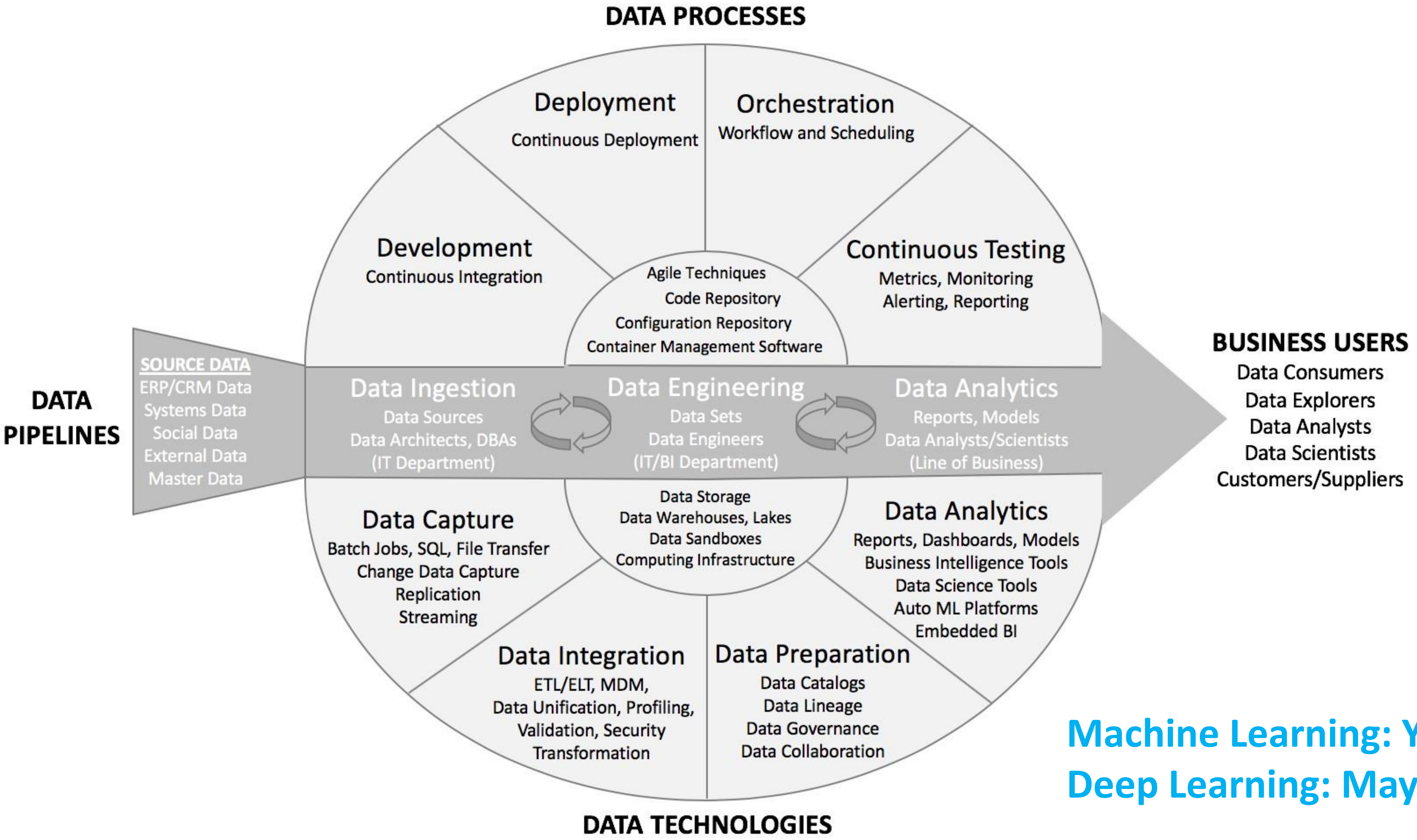
(Shahnas & Yuen, Constraints on Geophysical Parameters Using Machine Learning Algorithms, JGR 2018)





Scarcity of Data

Enterprise Computing versus Scientific Computing



Machine Learning: Yes
Deep Learning: Maybe

How do we educate our students in this new age?

Both graduate and undergraduate students

In this talk, I focus on undergraduate education from Interdisciplinary Undergraduate Program on Computational Sciences (계산과학 연합전공)

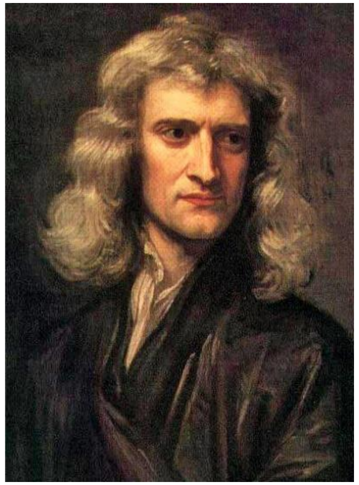
Currently 22 enrolled with about 50 taking courses (mostly from Physics, Mathematics, Earth Sciences and College of Liberal Studies)

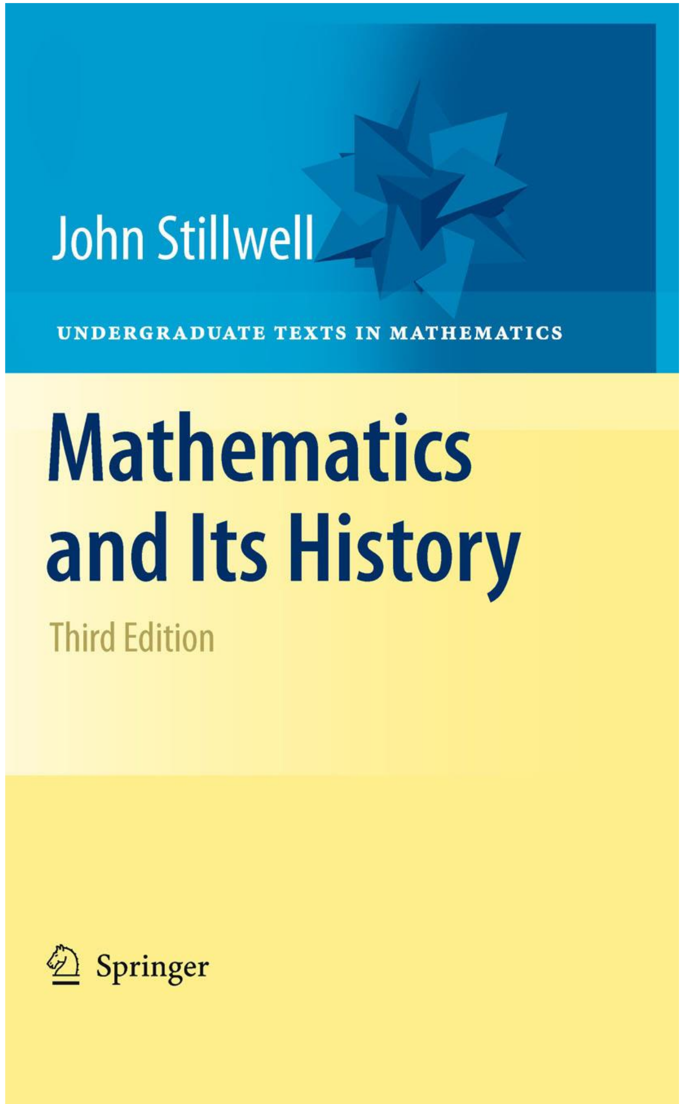
Curriculum

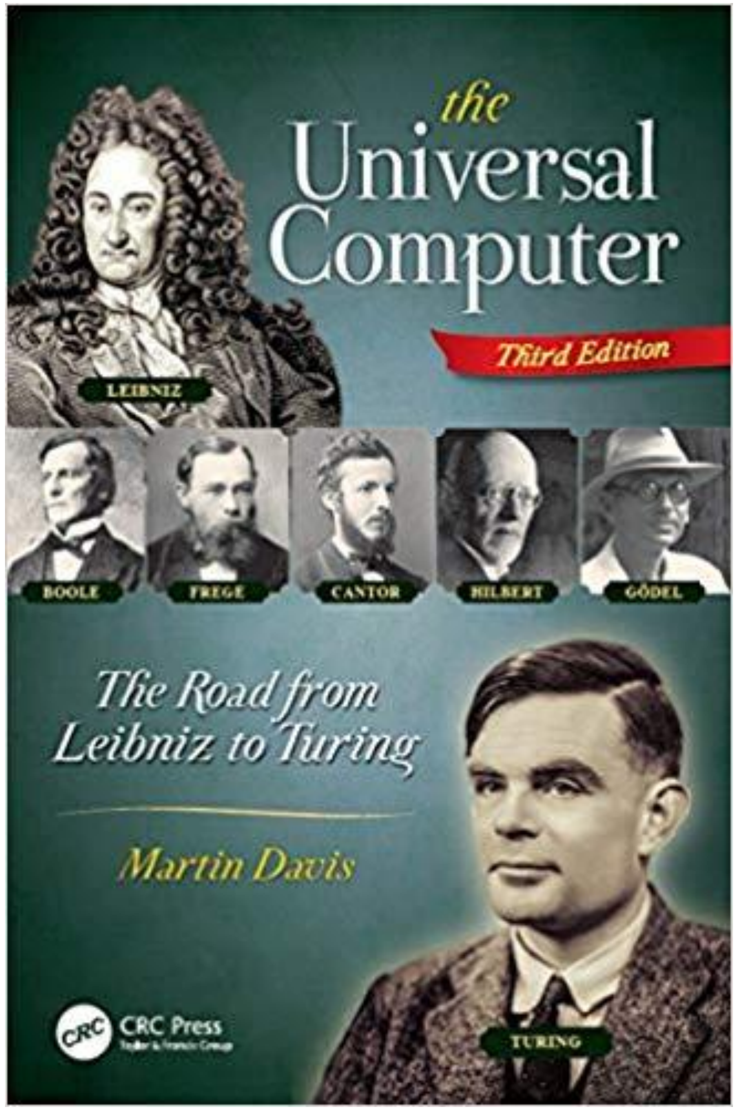
학년	1학기 교과목 및 학점	2학기 교과목 및 학점
2학년	• 계산과학의 이해 (Understanding Computational Sciences) • 계산과학의 기초와 역사 (History and Foundation of Computational Sciences) • 계산과학이론 및 실습 1 (Theory and Practices of Computational Sciences 1)	• 계산과학이론 및 실습 2 (Theory and Practices of Computational Sciences 1)
3학년	• 데이터과학 (Data Sciences)	• 과학계산개론 (Introduction to Scientific Computing)
4학년		• 계산과학 주제연구 (Special Topics On Computational Sciences) • 계산과학 종합설계 (Capstone Research on Computational Sciences)

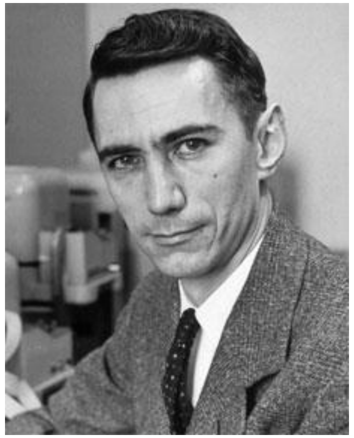
6 Prerequisites and 2 Electives

Foundation and History of Computational Sciences

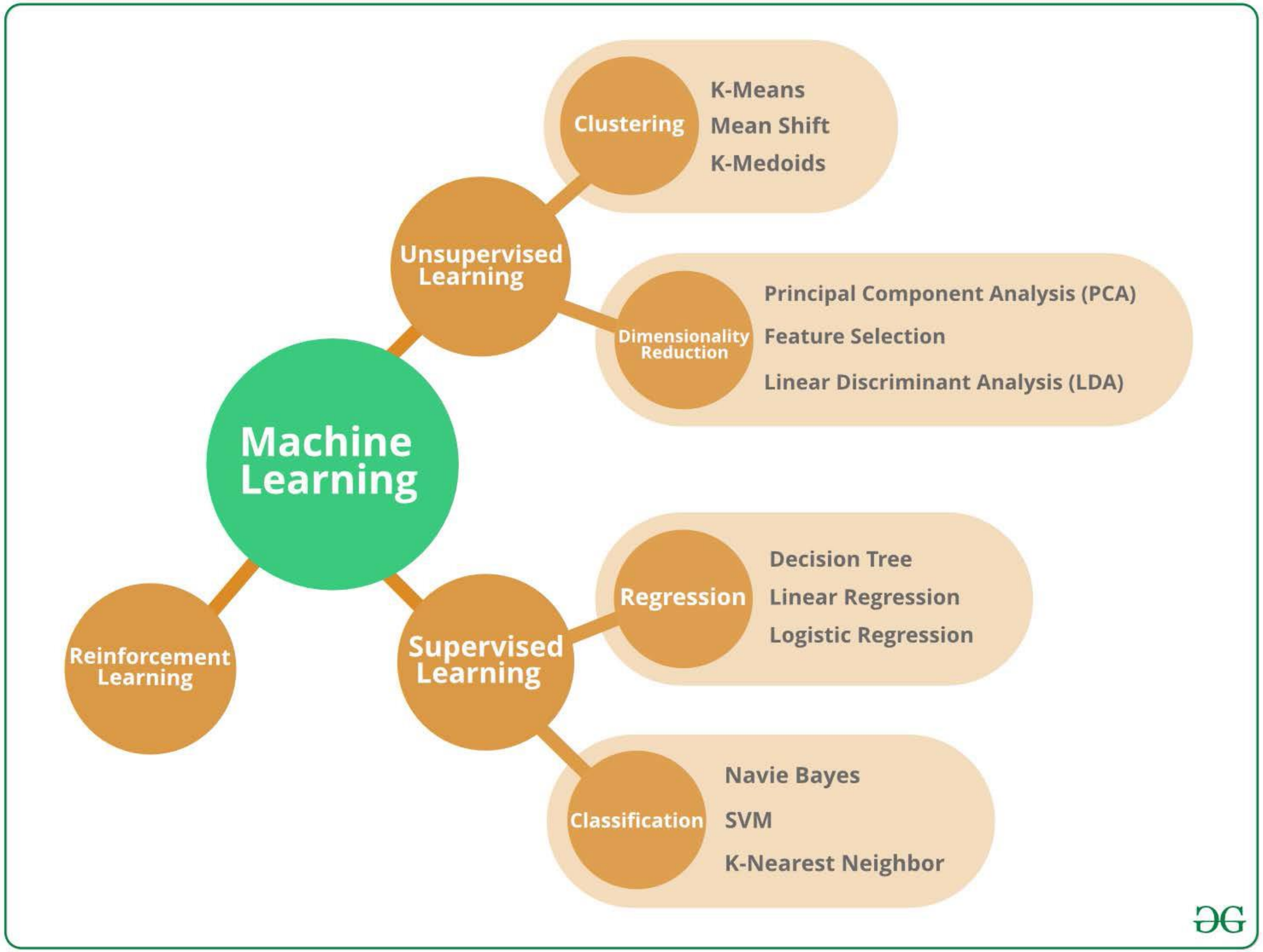








Leibniz
Boole
Frege
Cantor
Russell
Hilbert
Gödel
Turing
Shannon

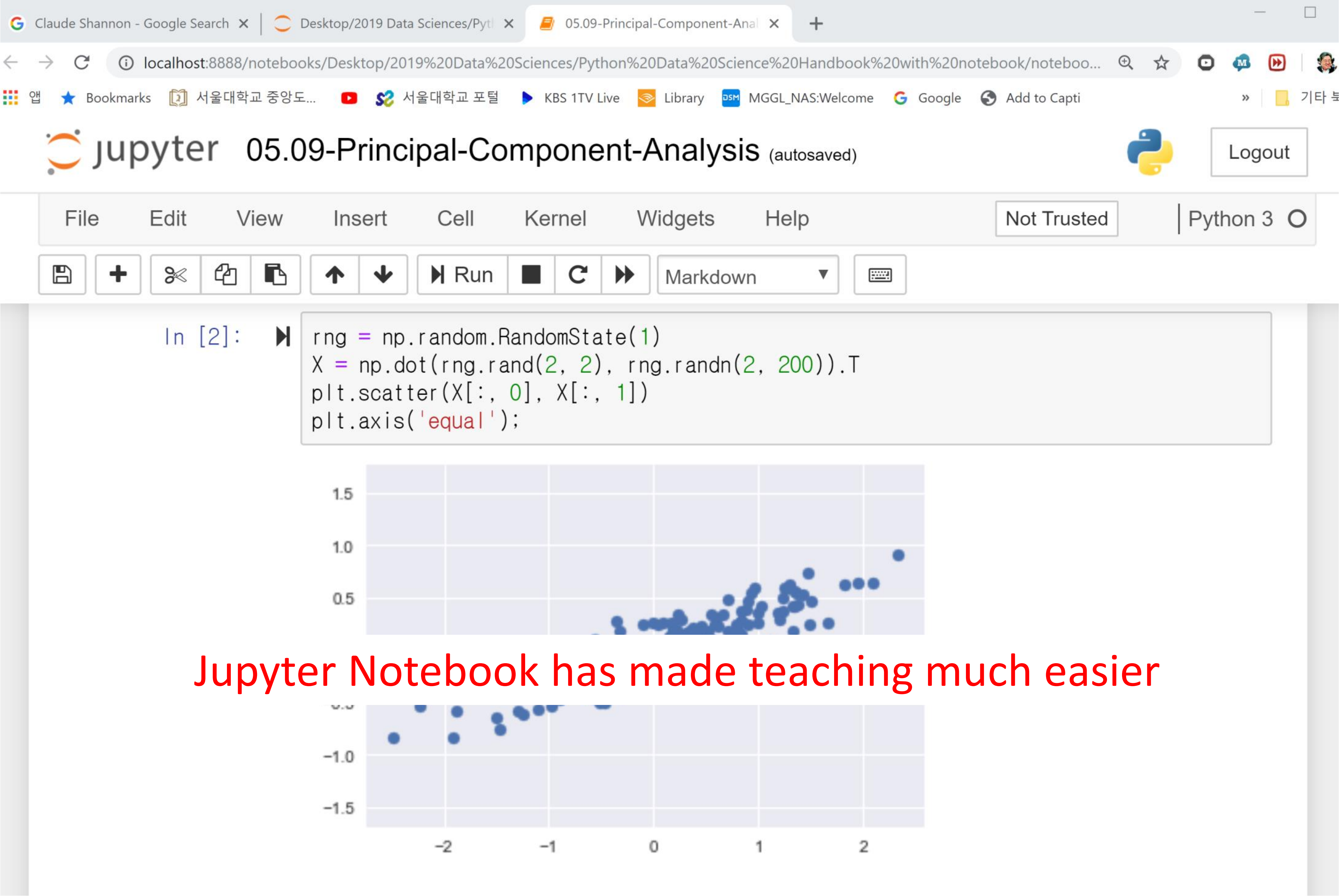


With special emphasis on the principle and applications of

- **Linear Algebra**
- **Probability**
- **Optimization**

as well as becoming familiar with programming tools such as

- **Python (numpy, pandas, matplotlib)**
- **C/C++**

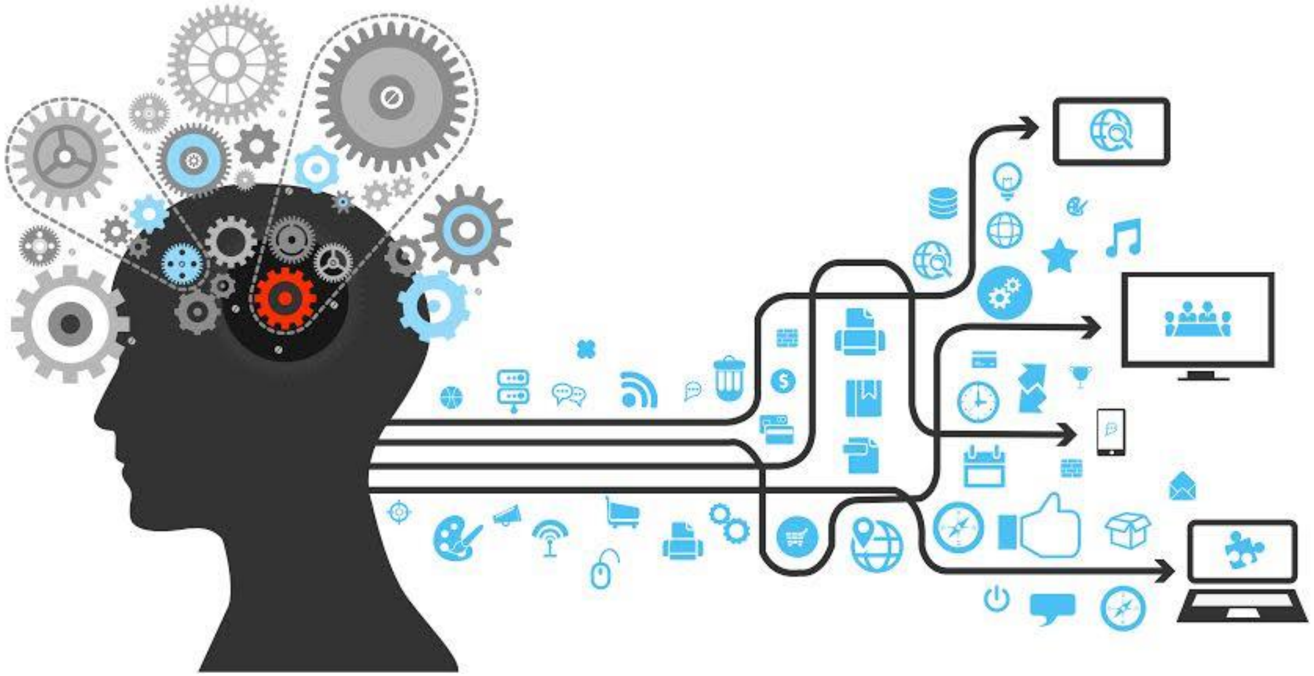


Jupyter Notebook has made teaching much easier

Special Topic: Designing Software Apps for People with Disability

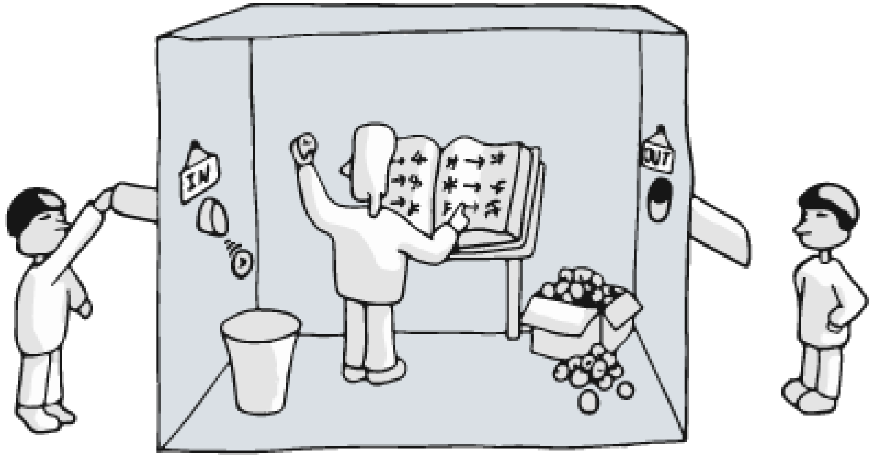
AT EDUCOM 2017 at UAE

AI (Artificial Intelligence)
Weak vs **Strong**

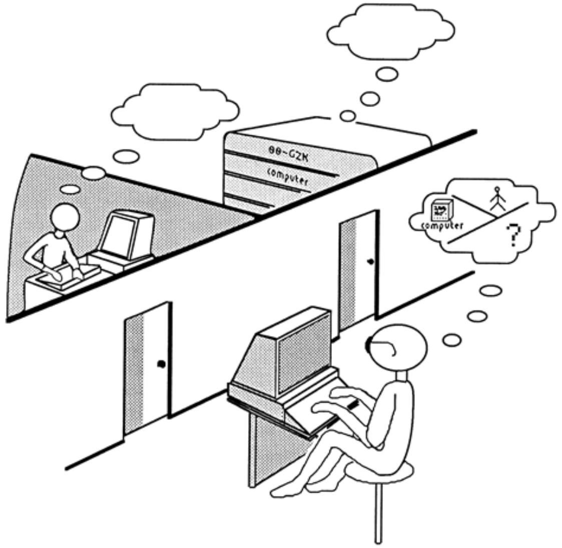


I think emergence of AI has led to question about our self such as "are we a machine?"

John Searle



Chinese Room Argument



Turing Test

Subjectivity(주관성) vs Objectivity(객관성)

Epistemic sense
신윤복은 화가이다. (objective)
신윤복은 우리나라에서 제일 유명한 화가이다. (subjective)

Ontological sense
산, 암석, 분자/원자, 세포, 인간 (objective, 관찰자와 무관)
돈, 명예, 결혼, 국가, 정보,계산 (subjective, 관찰자에 따라 다름)

Syntactical vs Semantic

There is no Turing Machine in Nature. It is just our interpretation.

Future direction

- *We live in an exciting time in human history (and even in Earth's history).*

- Students need to be exposed to workstations and supercomputing at early stage.
- In addition to computer literacy, we need to emphasize domain knowledge and underlying complexity. (importance of observations and experiments to overcome lack of data)
- Identify important societal problems and motivate students to come up with solution using science and technology.
- Get new dedicated faculty member for computational sciences.
- Make education truly interdisciplinary (perhaps transdisciplinary).

Thank You Very Much for Your Attention