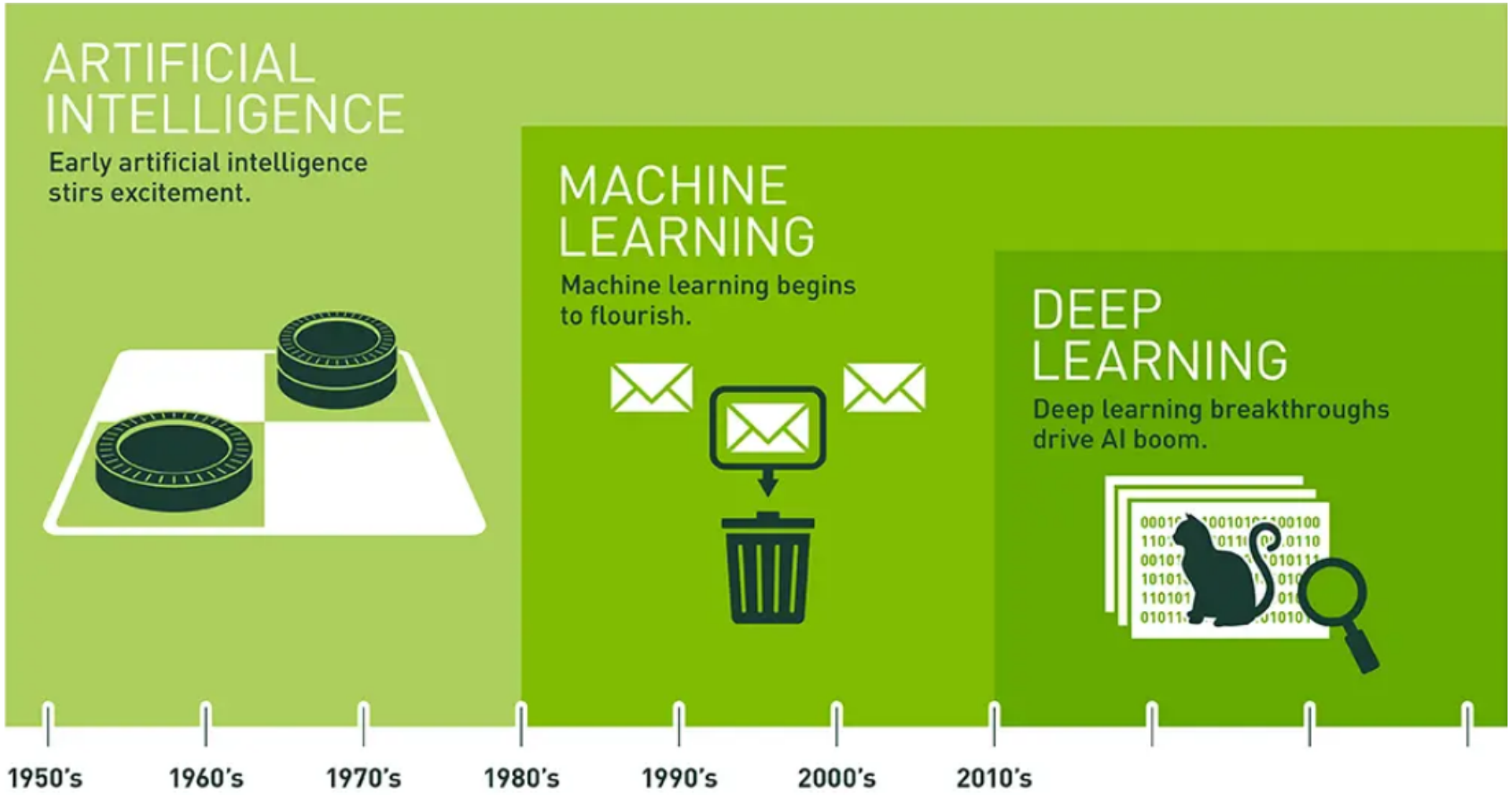
🎯 강의목표

김수환

https://www.soohwan.kim

인공지능? 

AI vs. ML vs. DL



https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/

AI vs. ML vs. DL

- AI
  - Algorithms: Problem Solving
  - Data Structures: Queue, Stack, Tree, ...
  - Search, Logic, Reasoning, Knowledge Representation
- ML
  - Supervised Learning: Regression, Classification
    - Model + Data + Parameter Optimization
  - Unsupervised Learning: Clustering
  - Reinforcement Learning
  - Handcrafted Features
- DL
  - Model: deep neural networks with billions of parameters
  - Data: Big Data
  - Optimization Techniques: batch normalization, pruning
  - Learned Features

Supervised Learning in ML and DL

1. 학습데이터를 만든다. (**입력값** vs. **출력값**)
2. 모델을 만든다 (**파라미터**가 포함된 **함수**):
3. 학습데이터를 이용하여 모델을 학습한다. (파라미터를 최적화한다)
4. 테스트데이터의 입력값을 함수에 넣어 출력값을 구한다.

$$\mathbf{y} = f(\mathbf{x}; \mathbf{w})$$

생각해 보기

다음에 나올 수는?

$$0, 2, 4, 6, 8, ?$$

다음에 나올 수는?

$$0, 2, 4, 6, 8, 10$$

다음에 나올 수는?

$$0, 2, 4, 6, 8, 10$$

$$1, 3, 5, 7, 9, ?$$

다음에 나올 수는?

$$0, 2, 4, 6, 8, 10$$

$$1, 3, 5, 7, 9, 11$$

다음에 나올 수는?

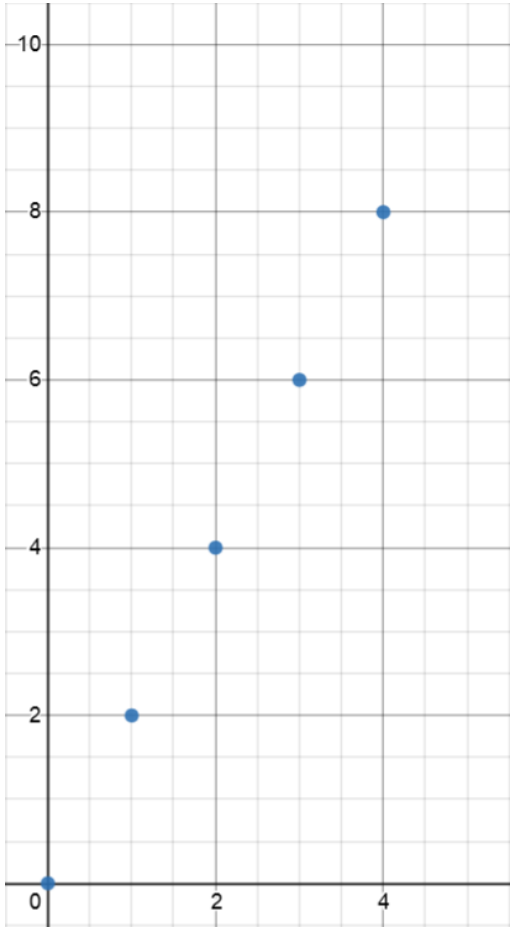
$$0, 2, 4, 6, 8, 10$$

$$1, 3, 5, 7, 9, 11$$

왜?

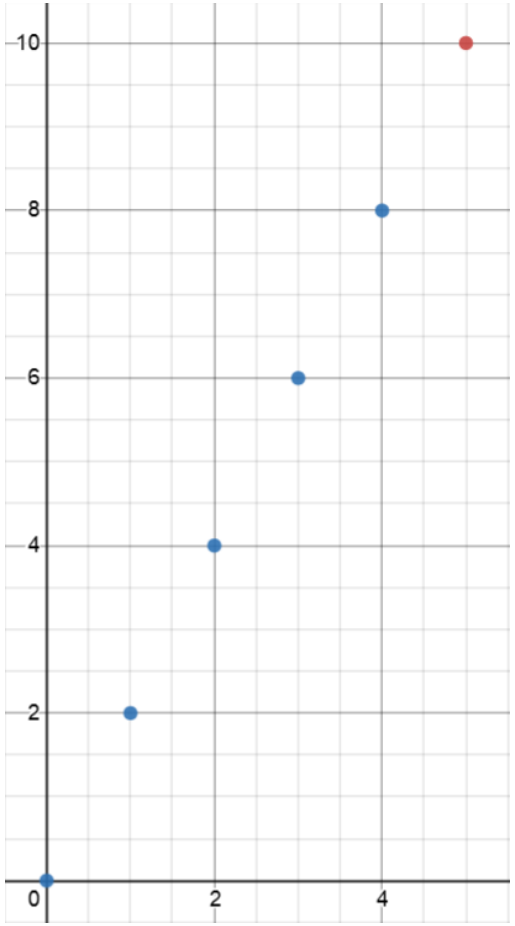
Human Intelligence (인간지능)

입력	출력
0	0
1	2
2	4
3	6
4	8
5	?



Human Intelligence (인간지능)

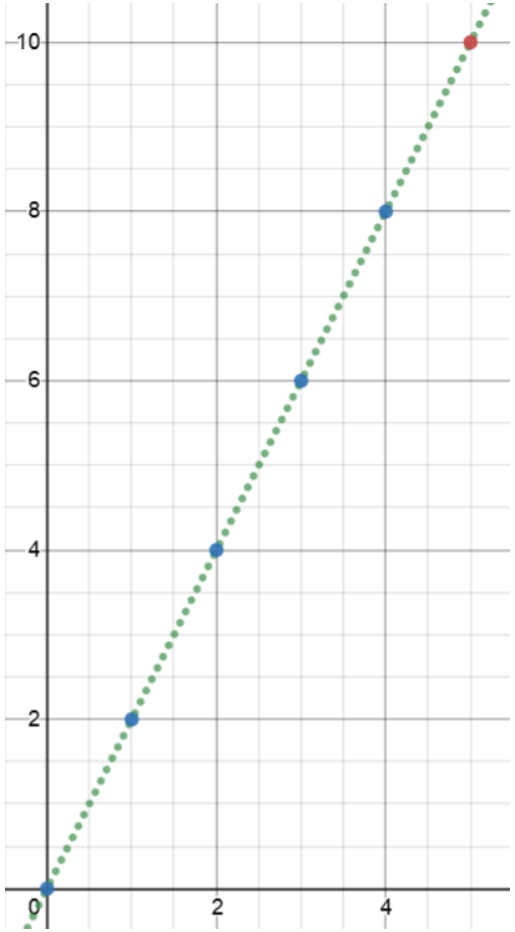
입력	출력
0	0
1	2
2	4
3	6
4	8
5	10



Human Intelligence (인간지능)

입력	출력
0	0
1	2
2	4
3	6
4	8
5	10

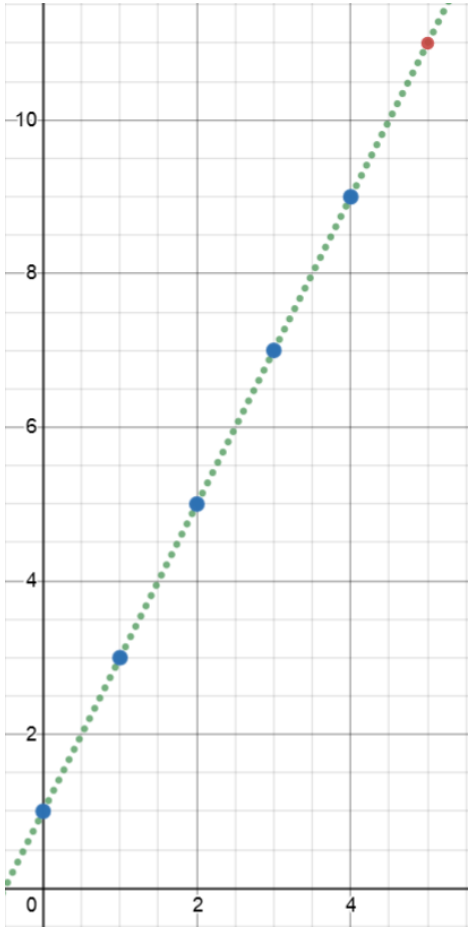
$$y = 2x$$



Human Intelligence (인간지능)

입력	출력
0	1
1	3
2	5
3	7
4	9
5	11

$$y = 2x + 1$$



다음에 나올 수는?

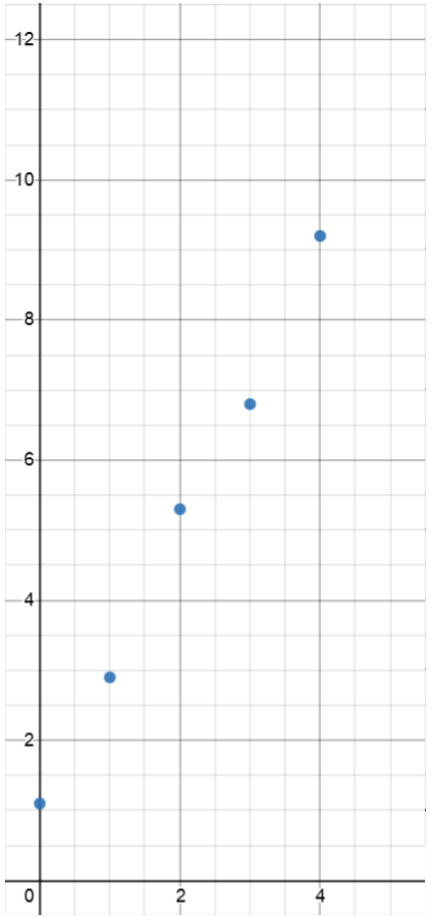
$$0, 2, 4, 6, 8, 10$$

$$1, 3, 5, 7, 9, 11$$

$$1.1, 2.9, 5.3, 6.8, 9.2, ?$$

Machine Intelligence (인공지능)

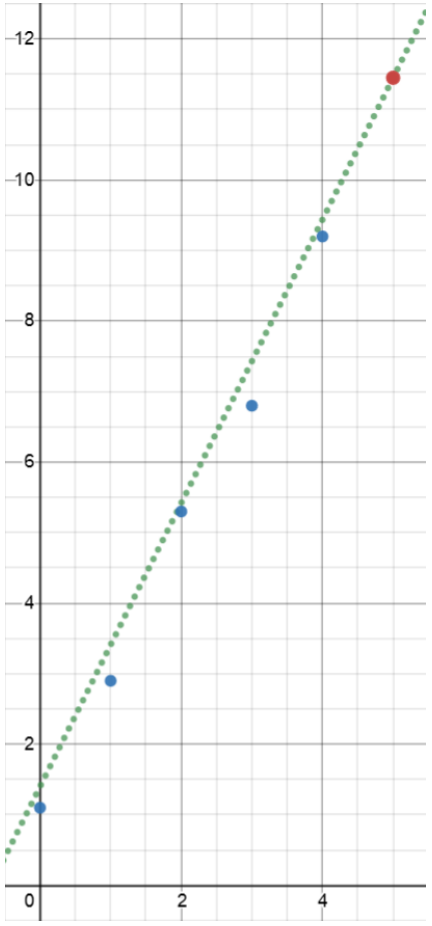
입력	출력
0	1.1
1	2.9
2	5.3
3	6.8
4	9.2
5	?



Machine Intelligence (인공지능)

입력	출력
0	1.1
1	2.9
2	5.3
3	6.8
4	9.2
5	11.45

$$y = 2.1x + 1.04$$

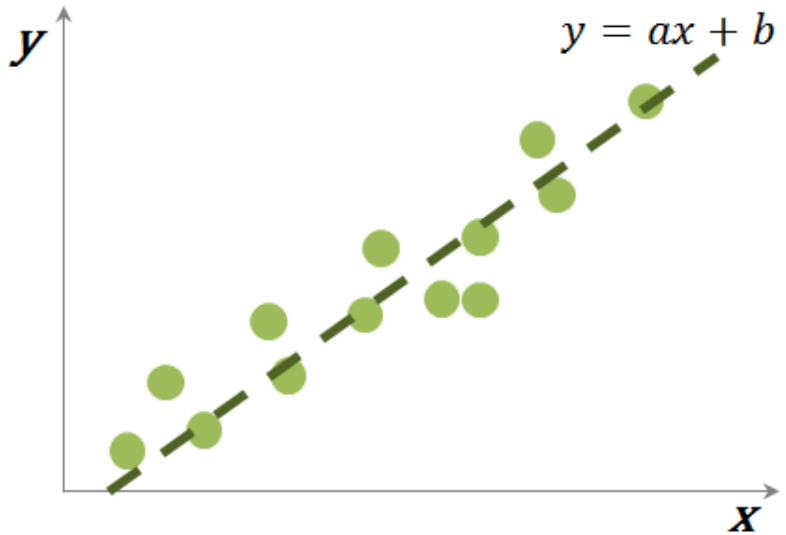


선형회귀

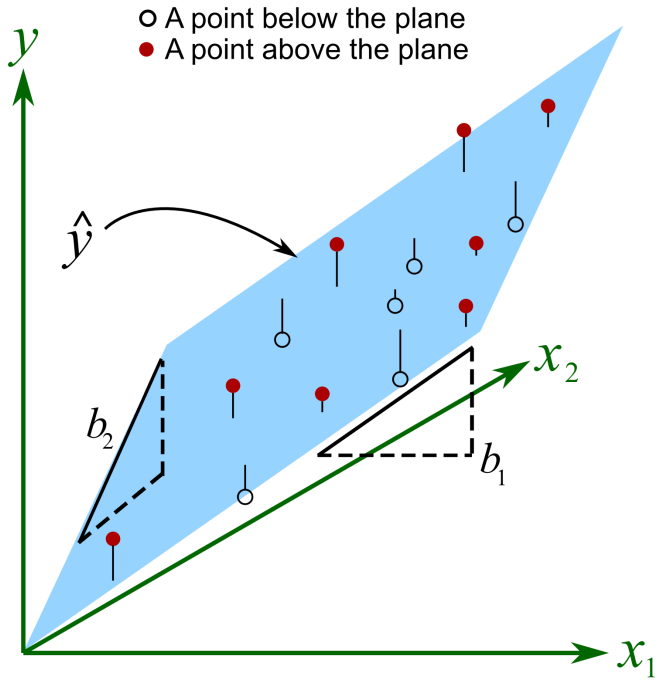
Linear Regression

선형회귀

- 1D


y
y = ax + b
x

- 2D


A point below the plane
A point above the plane
y
ŷ
x2
x1
b2
b1

풀이 ① 최소자승법 (Least Square Errors)

1. Data

$$D = \{(\mathbf{x}_i, y_i)^\top \mid \mathbf{x}_i \in \mathbb{R}^d, y \in \mathbb{R}\}_{i=1}^n \quad \Rightarrow \quad \mathbf{X} \in \mathbb{R}^{n \times d}, \mathbf{y} \in \mathbb{R}^n$$

2. Model

$$y = \mathbf{w}^\top \mathbf{x}$$

3. Error

$$e = (\mathbf{y} - \mathbf{X}\mathbf{w})^\top (\mathbf{y} - \mathbf{X}\mathbf{w})$$

4. Partial Derivative

$$\frac{\partial e}{\partial \mathbf{w}} = 0 \quad \Rightarrow \quad \mathbf{w} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$$

풀이 ② 의사역행렬 (Pseudo Inverse)

1. Data

$$D = \{(\mathbf{x}_i, y_i)^\top \mid \mathbf{x}_i \in \mathbb{R}^d, y \in \mathbb{R}\}_{i=1}^n \quad \Rightarrow \quad \mathbf{X} \in \mathbb{R}^{n \times d}, \mathbf{y} \in \mathbb{R}^n$$

2. Model

$$y = \mathbf{w}^\top \mathbf{x}$$

3. System of Equations

$$\mathbf{y} = \mathbf{X}\mathbf{w}$$

4. Pseudo Inverse

$$\mathbf{w} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$$

풀이 ③ 최대 우도 (Maximum Likelihood)

1. Data

$$D = \{(\mathbf{x}_i, y_i)^\top \mid \mathbf{x}_i \in \mathbb{R}^d, y \in \mathbb{R}\}_{i=1}^n \quad \Rightarrow \quad \mathbf{X} \in \mathbb{R}^{n \times d}, \mathbf{y} \in \mathbb{R}^n$$

2. Model

$$y = \mathbf{w}^\top \mathbf{x} + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \sigma^2)$$

3. Likelihood

$$P(D \mid \mathbf{w}) = \prod_{i=1}^{n} \mathcal{N}(y_i; \mathbf{w}^\top \mathbf{x}_i, \sigma^2)$$

4. Maximum Likelihood

$$\frac{\partial}{\partial \mathbf{w}}(-\log P) = 0 \quad \Rightarrow \quad \mathbf{w} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$$

선형회귀의 3가지 풀이법

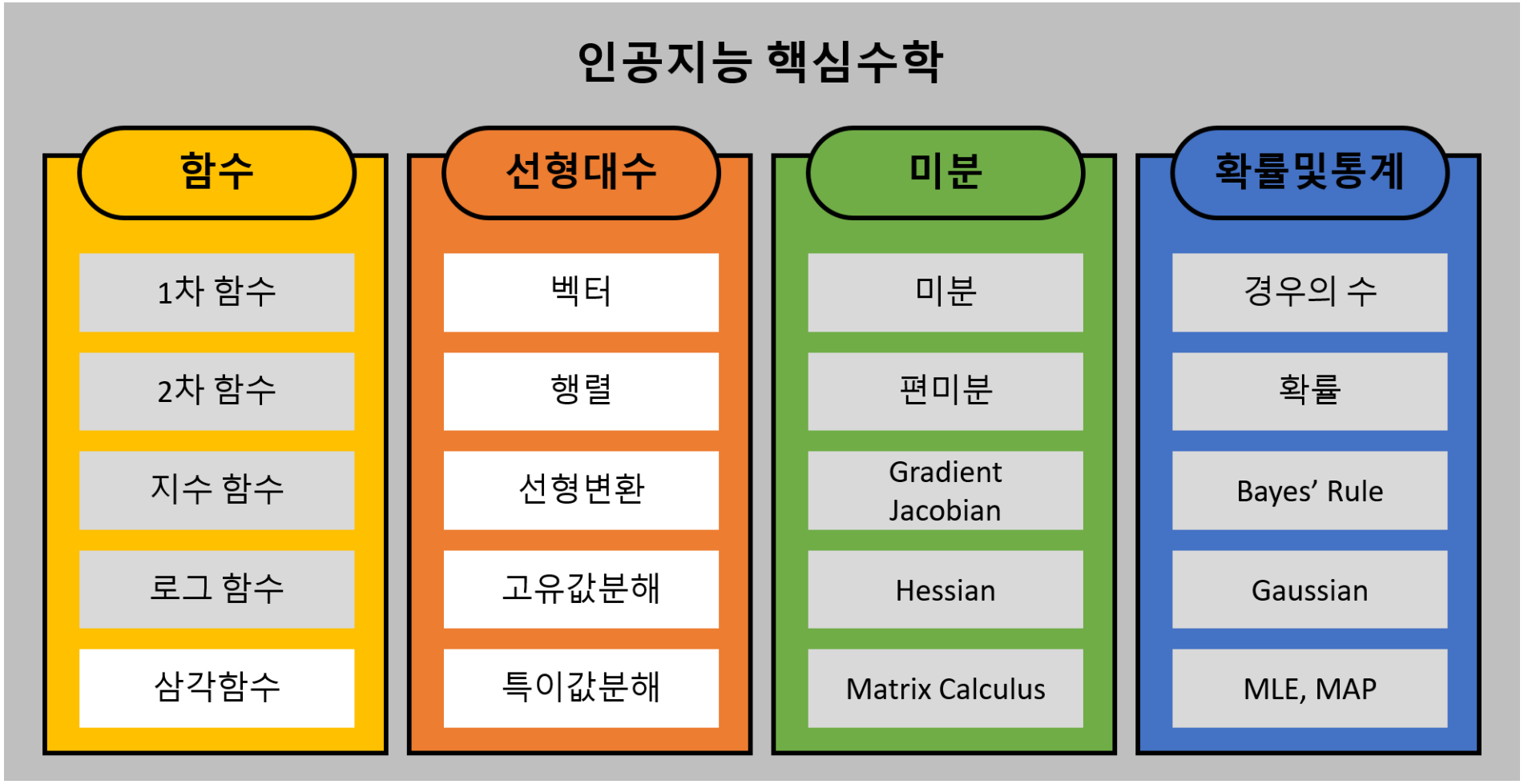
http://m.ppomppu.co.kr/new/bbs_view.php?id=help&no=1257492

선형회귀의 3가지 풀이법

https://giphy.com/gifs/afterthis-oleeee-3oFyDoJts3s3rIb6O4

👀 앞으로 배울 내용

Big Picture


인공지능 핵심수학
함수
1차 함수
2차 함수
지수 함수
로그 함수
삼각함수
선형대수
벡터
행렬
선형변환
고유값분해
특이값분해
미분
미분
편미분
Gradient
Jacobian
Hessian
Matrix Calculus
확률및통계
경우의 수
확률
Bayes' Rule
Gaussian
MLE, MAP

The Unit Cube

Unit Cube (단위정육면체)