지능 시스템
Intelligent Systems

Lecture 0 – Introduction to Artificial Intelligence and Deep Learning

교수 소개

온성권 중위(진)

- 학사 146기
- University of Oxford 학.석사 졸업

Machine Learning 전공

- Reinforcement Learning (강화학습) 연구
  - Oxford Robotics Institute (ORI), Applied Reinforcement Learning in Robotics 연구

연락처:

- 단재관 336호, (043)290-6584
- 개인 폰번호: 010-5791-1038
- E-mail: 21-10612@mnd.go.kr

일정 (1/2)

세부 강의계획 (월1-2, 화3-4)

주차	월	주	일	교육 내용	수업방식
1	8	5	28 29	Introduction to Reinforcement Learning, MDP, and Open AI Gym	강의식
2	9	1	4 5	* **휴강**(해외항법/ 9. 5. ~ 9.)	-
3		2	11 12	* **휴강**(국군의날 훈련/ 9. 11. ~ 26.)	-
4		3	18 19		
5		4	25 26		
6	10	1	2	Bellman Equation * **임시 공휴일의 경우, 보충수업**	강의식
			3	* **휴강**(개천절 휴일)	-
7		2	9	* **휴강**(한글날 휴일)	-
			10	Value-based Algorithms (MC, SARSA, Q-Learning)	강의식
8		3	16 17	**중간시험(지필고사/ 비중 30%)**	

일정 (2/2)

중간고사: 필기시험 30%
기말고사 필기시험 30%
수시평가
- 문제풀이 과제 1회 or 2회 20%
- 코딩 리포트 1회 20%

9		4	23		
			24	DQN + variants (double, PER, duel, Rainbow)	강의식
10		5	30		
			31	* 휴강(무용기 결승/ 성무제)	-
11	11	1	6	Policy Gradients (REINFORCE)	
			7	Advantage Actor Critic (A2C)	
12		2	13	Proximal Policy Optimization (PPO)	
			14	Deep Deterministic Policy Gradient (DDPG & TD3)	
13		3	20	Soft Actor Critic (SAC)	강의식
			21	Model-based Algorithms (SimPLe, World Models)	
14		4	27	Tree Search and Alphago	
			28	과제 리뷰	
15	12	1	4	Python Revision	
			5	* 휴강(미래 항공우주 학술대회)	-
16		2	11	Implementation Practice	
			12	Implementation Practice	강의식
17		3	18	Implementation Practice	
			19	Implementation Practice	
18		4	25	* 휴강(크리스마스 휴일)	-
			26	Revision	강의식

참고 자료

공식 교재:

- 박성수, “수학으로 풀어보는 강화학습 원리와 알고리즘”

추가 추천 교재:

- 오가와 유타로, “Pytorch를 활용한 강화학습/심증강화학습 실전 입문”
- Sutton, Barto, “Reinforcement Learning”
- Ravichandiran, “Deep Reinforcement Learning with Python”
- 사이토 고키, “Deep Learning from Scratch(밑바닥부터 시작하는 딥러닝)”
- 최건호, “파이토치 첫걸음”

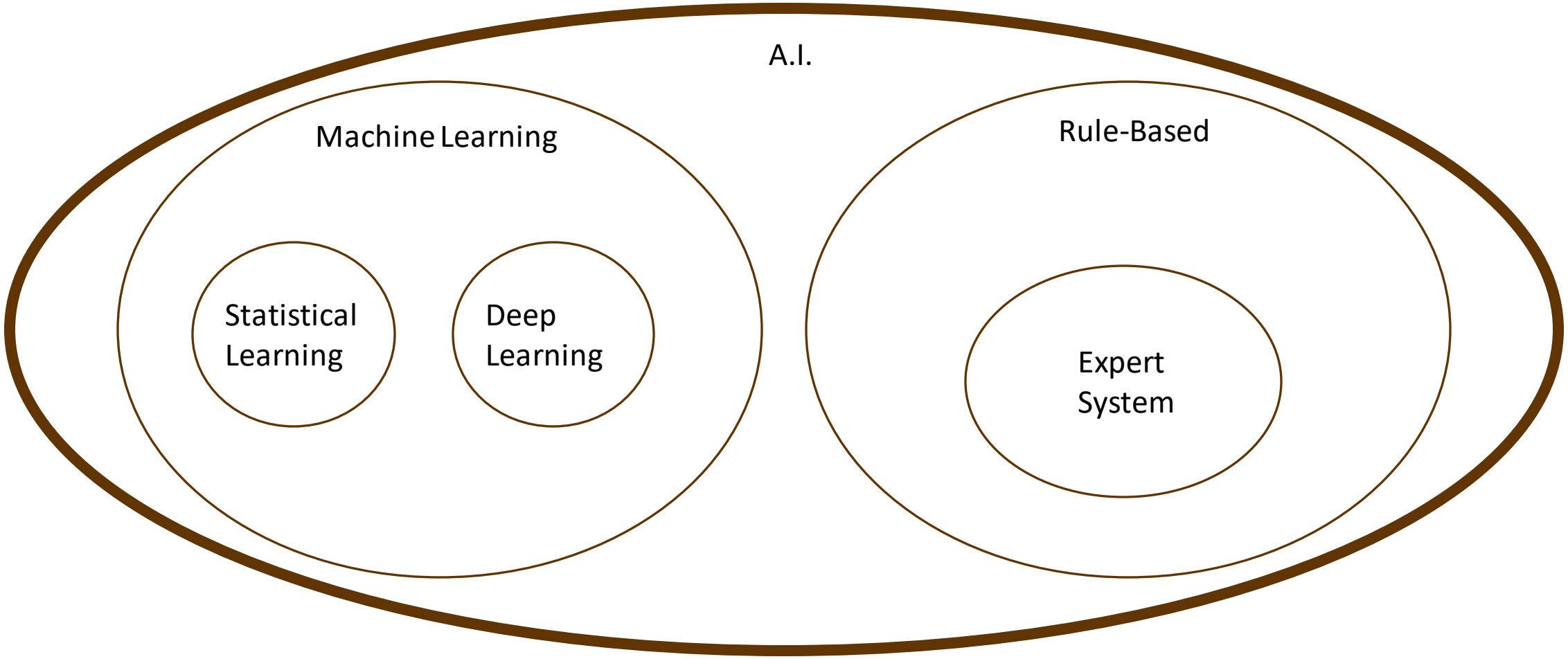
What is Artificial Intelligence?

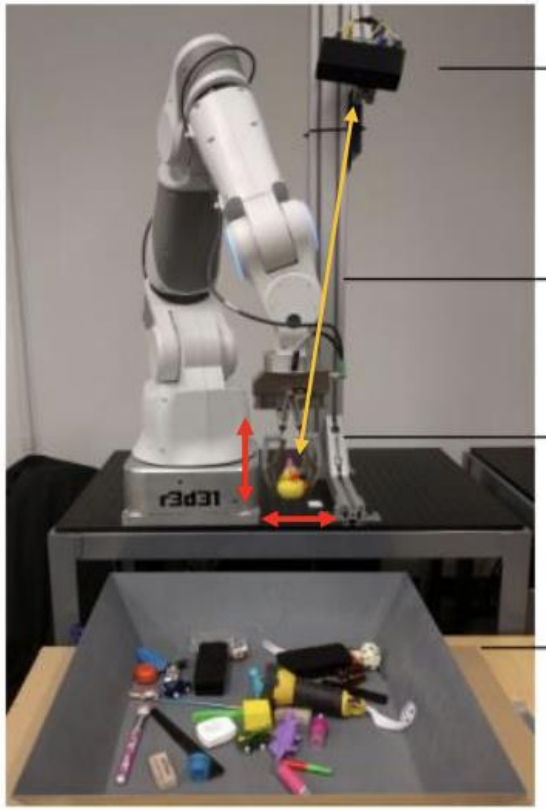
Wikipedia:
A.I. is the intelligence of machines or software, as opposed to the intelligence of human beings or animals.

Oxford Website:
The capacity of computers or other machines to exhibit or simulate intelligent behaviour

Chat GPT-3.5:
A.I. refers to the simulation of human-like intelligence in machines, enabling them to perform tasks that typically require human cognitive functions such as learning, reasoning, problem-solving, and decision-making.

Artificial Intelligence sub fields

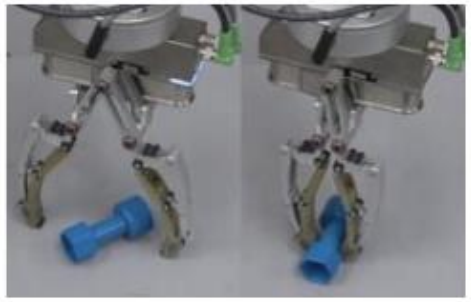
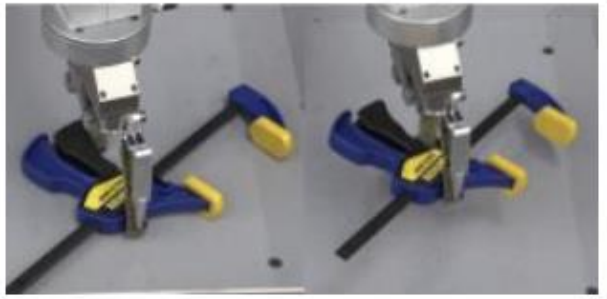
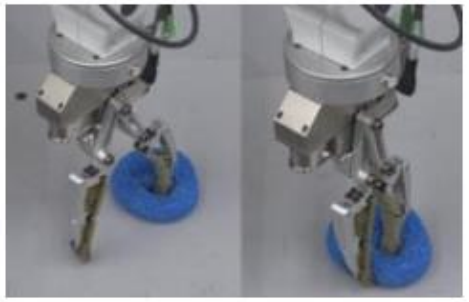




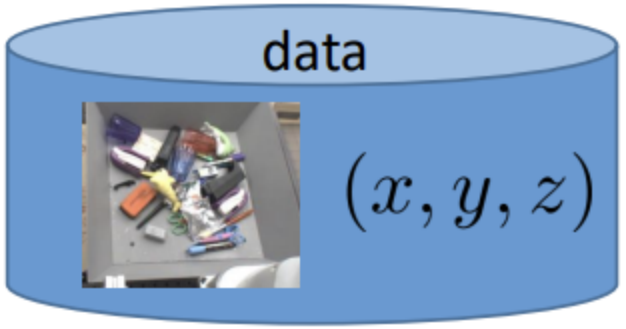
?

(x, y, z)

Method 1:
Understand the problem, design an appropriate solution

Method 2:
Set it up as a machine learning problem.



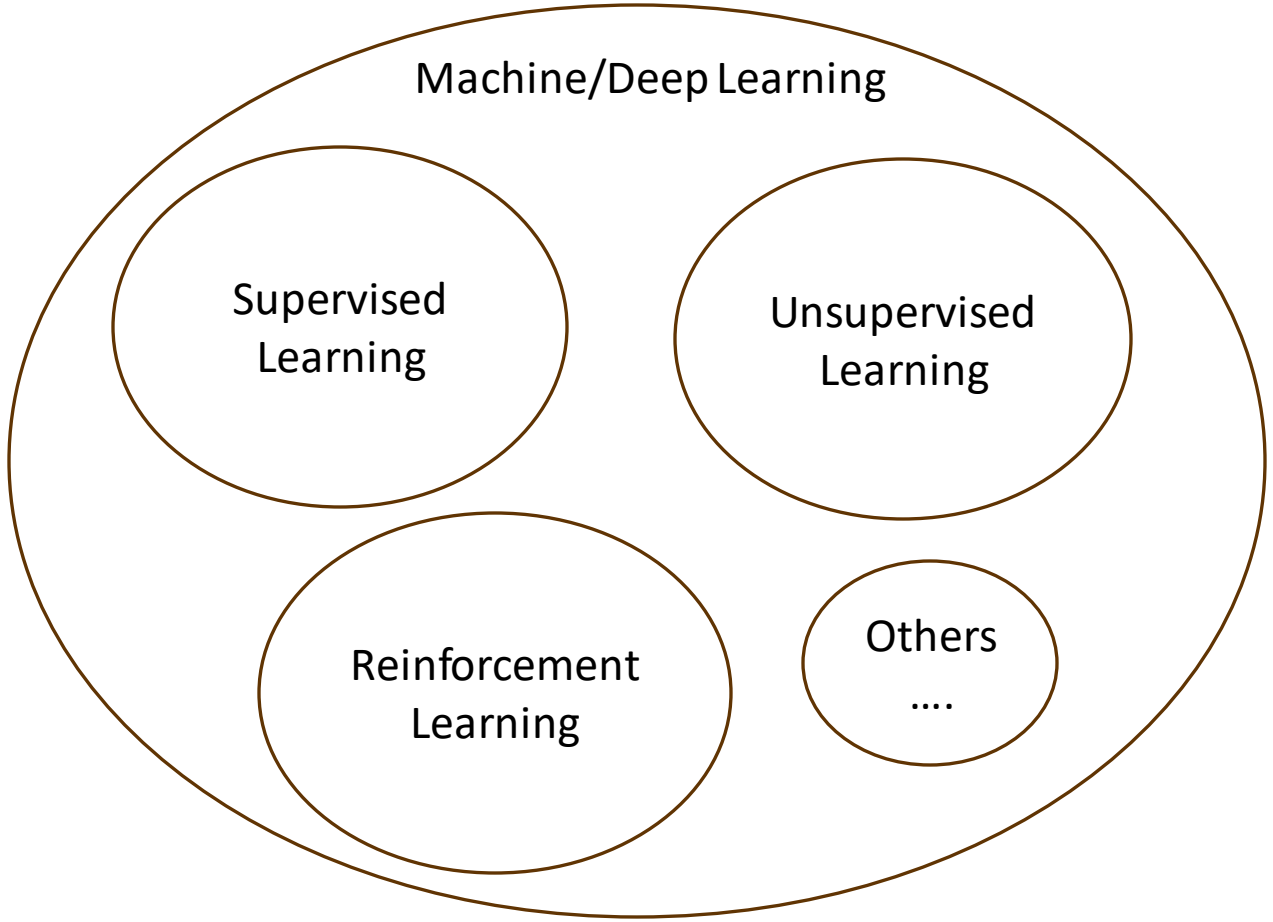
supervised learning

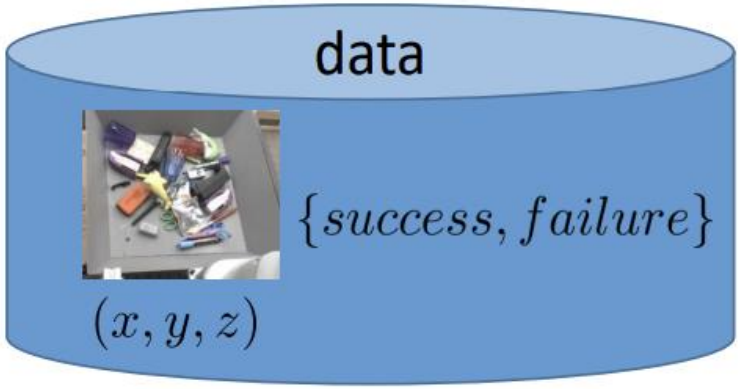
Machine/Deep Learning sub fields

Supervised Learning (SL):
Learning from data **with correct labels**.

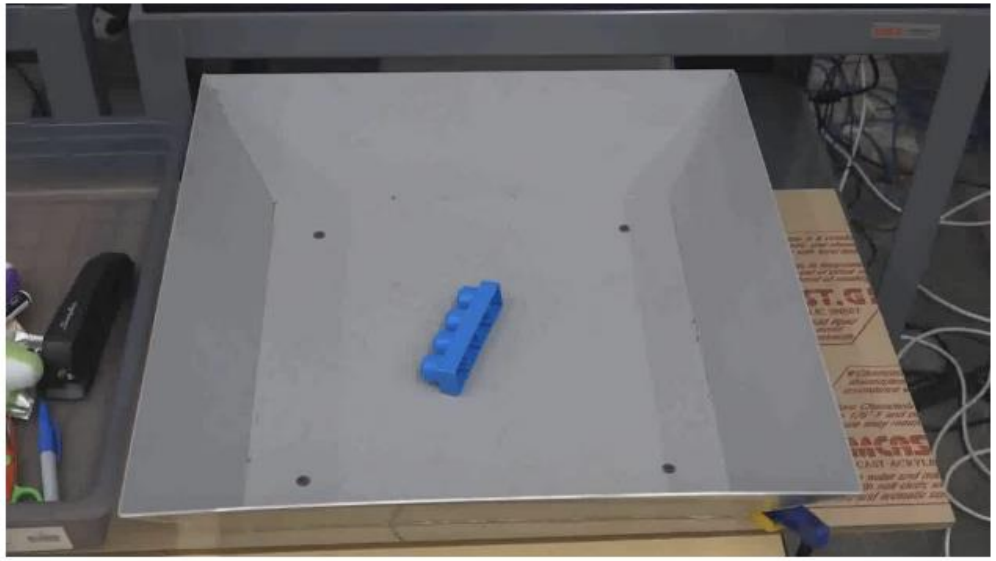
Unsupervised Learning (UL):
Learning from data **without labels**. The main objective is the clustering and grouping of data.

Reinforcement Learning (RL):
Learning through interactions with the environment, and constantly collecting **reward signals**, which act similar to labels in SL. However, the rewards are not absolutely trustworthy.



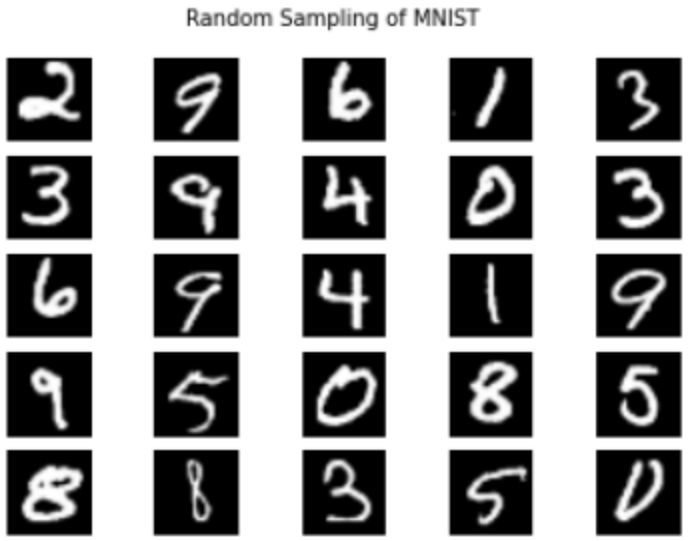


reinforcement learning



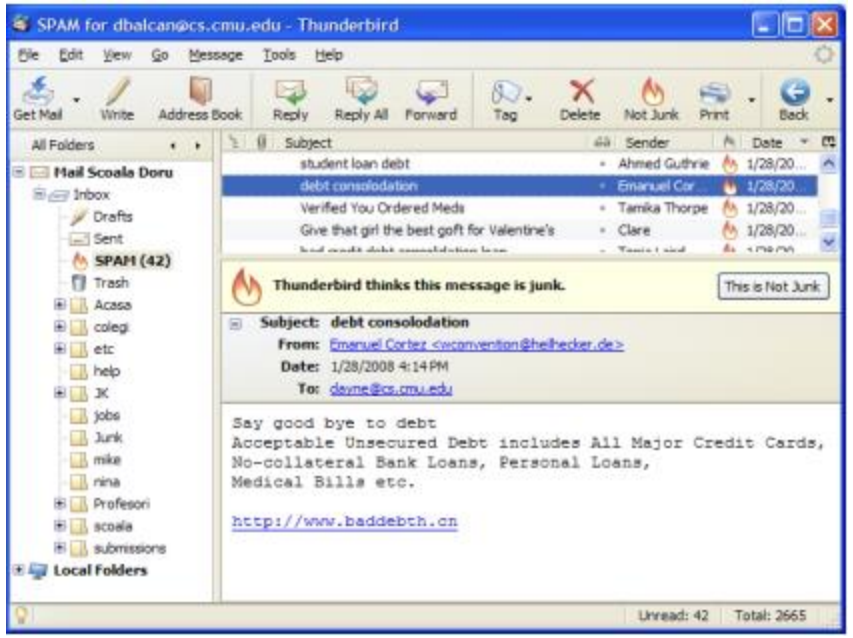
Supervised Learning (SL)

Hand-written digits recognition.



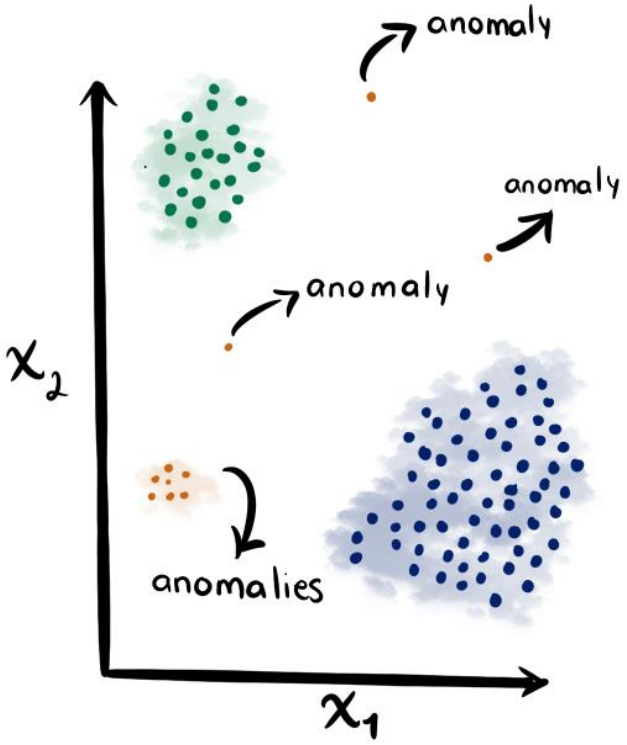
Face detection and recognition.

SPAM mail detection

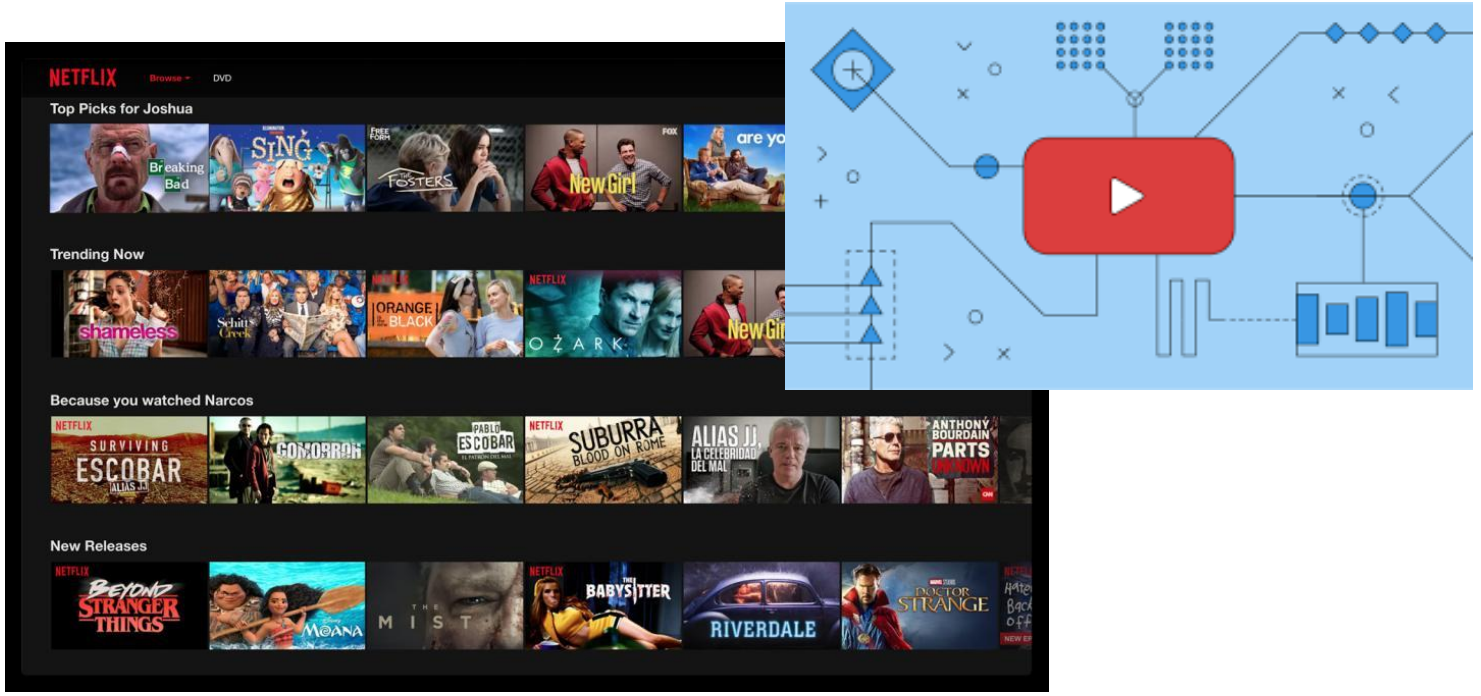


Unsupervised Learning (UL)

Anomaly Detection



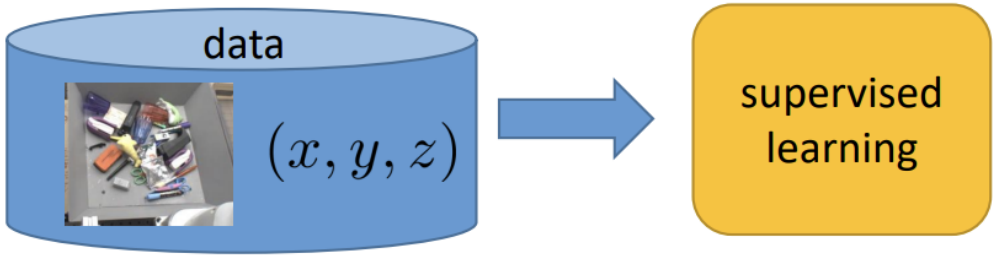
Personalized Recommender Engines



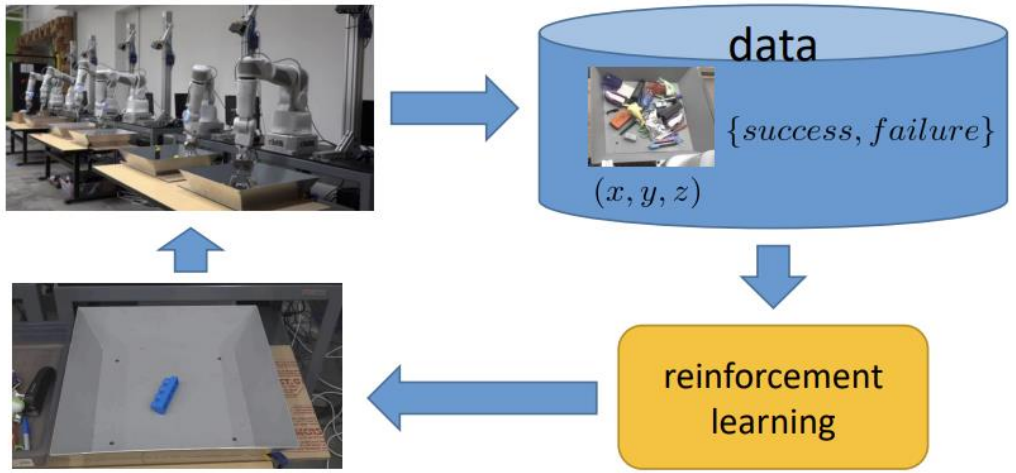
Reinforcement Learning (RL)
How is it different from SL?

Supervised Learning:
Given $\mathcal{D} = \{X_i, y_i\}$,
Learn how to predict y from X

Usually assumes
- i.i.d data
  (Independent, Identically Distributed)
- Known ground-truth outputs in training
  (A.K.A labels)

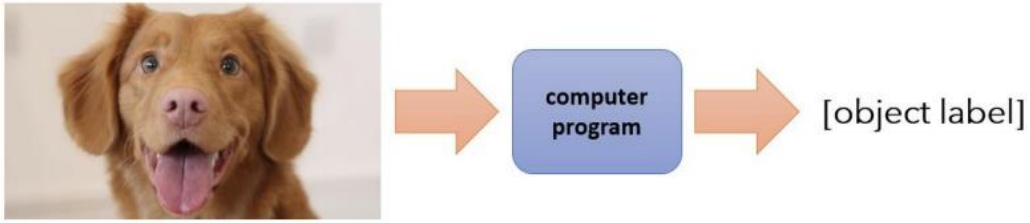


Reinforcement Learning:
- Data is not i.i.d. : previous outputs influence future inputs!
- Ground-truth labels are not given
  - We only know the reward!
- Data is constantly generated via interactions with the environment.



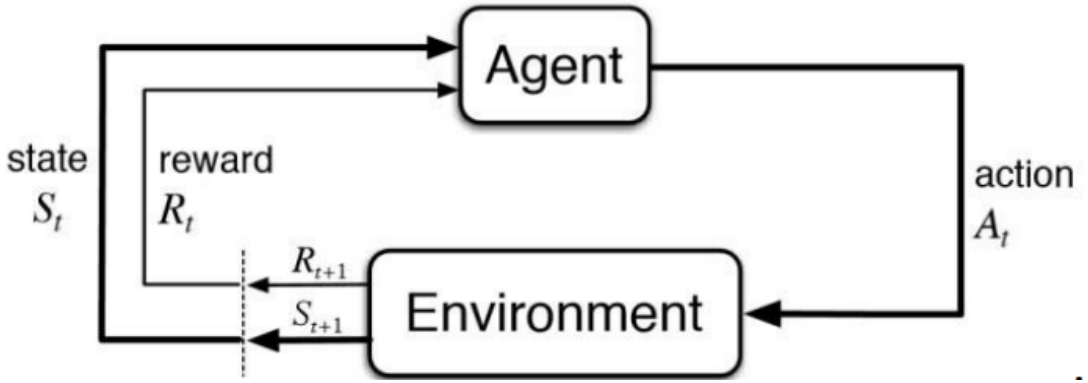
Reinforcement Learning (RL)
How is it different from SL?

Supervised Learning

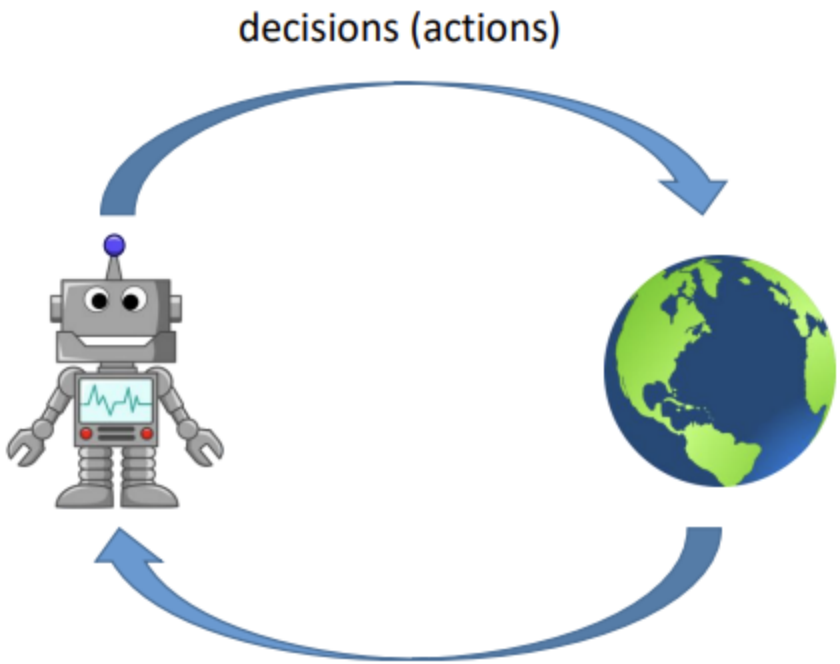


Input: X
Output: y
Data: $\mathcal{D} = \{X_i, y_i\}$
Goal: $f_\theta(X_i) \approx y_i$

Reinforcement Learning



Input: state, s_t at each time step
Output: action, a_t, at each time step
Data: $(s_1, a_1, r_1, \ldots, s_T, a_T, r_T)$
Goal: learn π_θ: $s_t \rightarrow a_t$
to maximize $\sum_t r_t$

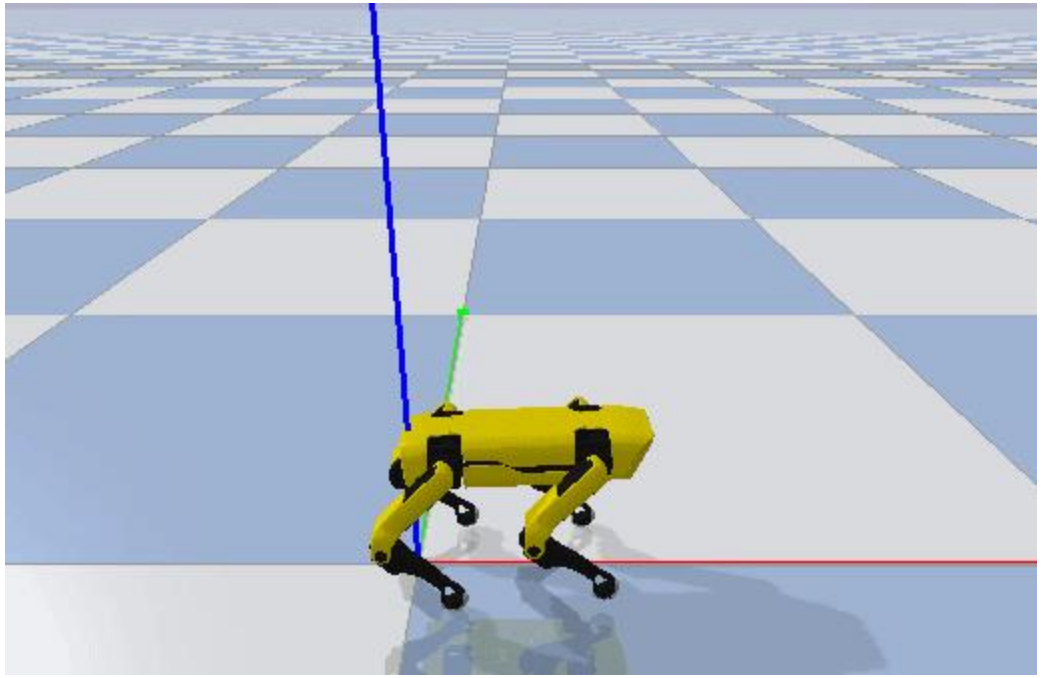


Actions: muscle contractions
Observations: sight, smell
Rewards: food

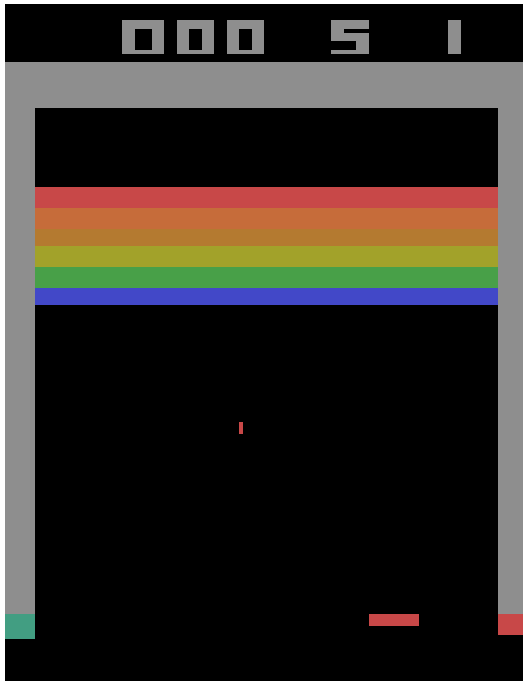
Actions: move
Observations: a new state of the play

Actions: Controls of F-16 (throttle, rudder, etc….)
Observations: a new coordinate of the fighter, distance to the adversary, bullets left, etc…

Complex Physical Tasks

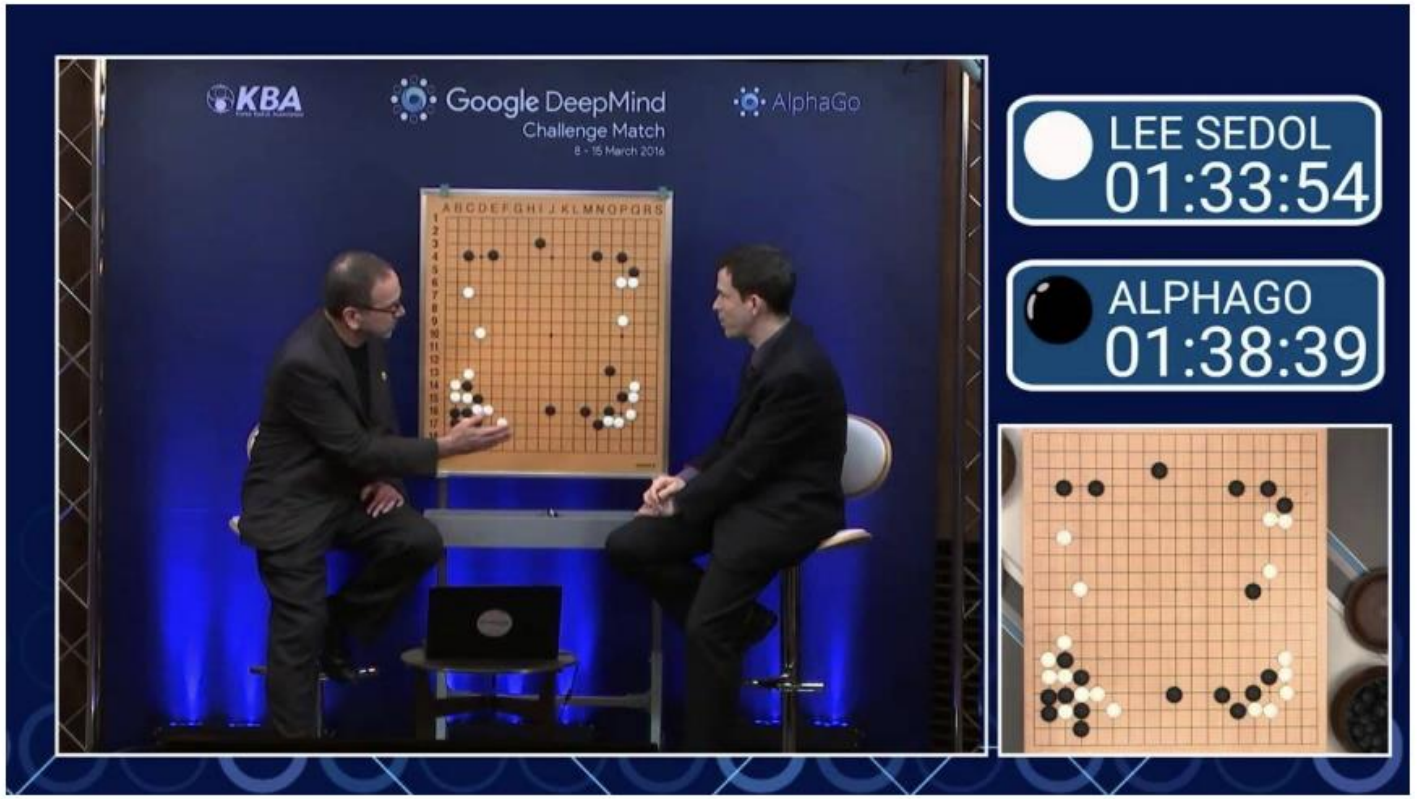


Discover optimum solutions!



Discover optimum solutions!

"Move 37" in Lee Sedol Alphago match: Reinforcement Learning "Discovers" a move that surprises everyone



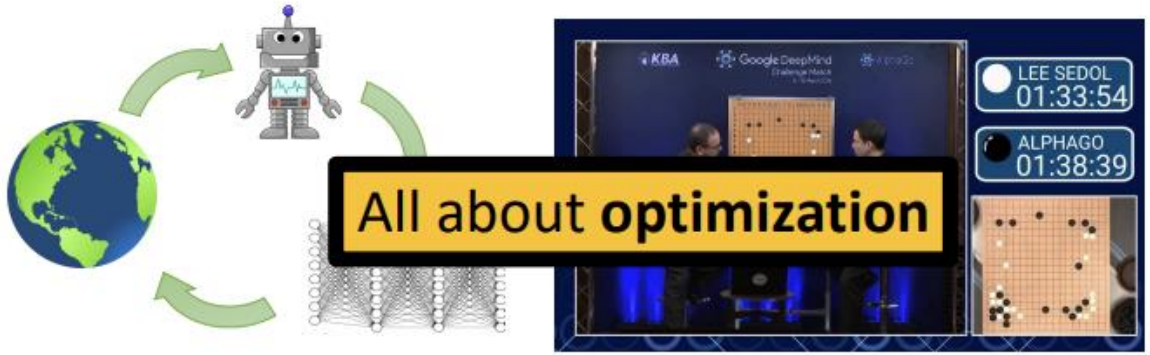
Data-driven AI VS RL

Data-Driven AI

+ learns about the real world from data

- doesn't try to do **better** than the data

Reinforcement Learning



+ optimizes a goal with emergent behavior

- but need to figure out how to use at scale!