

인공지능의 개요와 머신러닝의 학습방법

승실대학교 베어드교양대학
서유환 교수
yhsuh@ssu.ac.kr

학습 목표

- ▶ 인공지능의 정의와 분류를 설명할 수 있다
- ▶ 인공지능의 용어간 관계를 설명할 수 있다.
- ▶ 머신러닝의 학습방법(지도/비지도/강화)에 대해 설명할 수 있다.

목차

- ▶ 인공지능의 정의와 분류
- ▶ 머신러닝의 학습방법
 - ▶ 지도학습
 - ▶ 비지도학습
 - ▶ 강화학습

지난시간 배운 내용

주	주제	온라인	오프라인
1	인공지능의 과거 현재와 미래	1. 강의 및 교과목 소개(공통, 핵심만) 2. 인공지능의 과거와 현재 3. 인공지능의 미래와 다양한 시선 4. 인공지능 개발환경 구축과 사용법(Anaconda/Colab)	1. 강의 및 교과목 소개(분반별, 자세히) 2. 다양한 인공지능 기술 경험하기 (자연어처리, 시각, 음성) 3. 인공지능 챗봇만들기(IBM 왓슨 어시스턴트)
2	공공데이터를 이용한 사회문제 발견과 해결책 모색	1. 빅데이터의 정의와 가치 2. 공공데이터 수집하기 3. 공공데이터로부터 새로운 인사이트 발견하기 - 행정구역별 인구 데이터와 공공의료기관 현황 데이터 분석	1. 서울시 CCTV설치 현황 분석하기 2. 서울시 범죄발생 현황 분석하기
3	인공지능의 개요 및 머신러닝을 이용한 예측	1. 인공지능의 정의와 분류 2. 인공지능 학습방법 이해하기 3. 인공지능 알고리즘 소개	1. 머신러닝을 이용한 이미지 식별(구글 티처블 머신) 2. 머신러닝을 이용한 보스톤 집값 예측
4	인공지능과 데이터 윤리	1. 데이터의 불완전성과 결함에 따른 예측 오류와 차별 2. 데이터 왜곡에 따른 분석과 예측 결과 비교	1. 데이터 편향성이 예측에 미치는 영향 (구글 티처블 머신) 2. 데이터 왜곡에 따른 예측 결과 비교 - 타이타닉호 생존자 예측
5	인공지능과 알고리즘 윤리	1. 알고리즘 기반 의사결정 시스템의 한계 2. 윤리가 적용된 인공지능 알고리즘	1. 알고리즘에 따른 예측 결과 비교 - 보스톤 집값 예측 - 폐암환자 생존 여부 예측
6	인공지능에 대한 다양한 이슈와 우리의 자세 고찰	1. 인공지능의 윤리적/법적 쟁점 (자율주행자동차, AI로봇, 트랜스 휴먼 등) 2. 인공지능시대 사회, 경제적 불평등 문제 3. 인공지능과 프라이버시 4. 인공지능의 윤리적 대응과 규제	1. 자율주행 자동차의 행동학습 시나리오 경험하기 2. 비윤리적 데이터 생성과 수집(웹 크롤링을 이용한 데이터 수집)
7		기말고사	

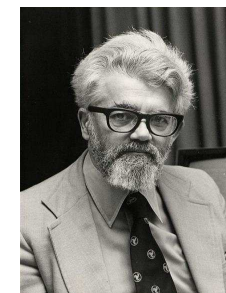
인공지능의 개요

인공지능의 정의

▶ 인공지능(AI, Artificial Intelligence)이란?

- ▶ 사람과 유사한 지능을 가지도록 인간의 학습능력, 추론능력, 지각능력, 자연어 이해능력 등을 컴퓨터 프로그램으로 실현하는 기술
- ▶ 문제 해결을 위해 상황을 인지하고, 파악하고 추론하여 답을 얻어내는 인간의 지능을 컴퓨터가 가질 수 있도록 실현한 기술
- ▶ 1956년 다트머스 학회에서 존 매커시(John McCarty) 교수가 처음 인공지능이라는 용어를 창안

"기계를 인간 행동의 지식에서와 같이 행동하게 만드는 것"



인공지능의 분류

인공지능의 용어간 관계

현재



좁은 인공지능
(Narrow AI)

- 스스로 사고해 문제를 해결할 수 있는 능력이 없는 컴퓨터 기반 AI
- 특정분야에 국한된 인공지능
- 특정 영역에서 인간보다 나은 성과
- AlphaGo, Watson, Siri

3 stages of AI



일반 인공지능
(General AI)

- 스스로 사고해 문제를 해결할 수 있는 능력을 가진 컴퓨터 기반 AI
- 인간 수준의 능력을 가진 인공지능
- 모든 분야에 적용될 수 있는 인공지능
- 자각력과 독립성을 갖춘



수퍼 인공지능
(Super AI)

- 인간보다 100~1000배 (IQ 1만~10만) 뛰어난 지능을 가진 AI
- 효율, 자기보존, 자원획득, 창의성 같은 원초적 욕구를 기반으로 끊임없이 자기 발전

Artificial Intelligence Simulates human intelligence

Narrow AI
Specific tasks

Machine Learning
Learns on past data

Deep Learning
A particular learning structure
- deep neural networks

General AI
~ Human

Super AI
>> Human

- 인공지능(Artificial Intelligence)
- 인간과 비슷하게 사고하는 컴퓨터의 지능을 구현하는 포괄적 개념

- 머신러닝(Artificial Intelligence)
- 데이터를 입력해 컴퓨터를 학습시키거나 스스로 배우게 해 인공지능 성능을 향상시키는 방법

- 딥러닝(Deep Learning)
- 머신러닝의 한 분야로 신경망(Neural Network)을 통해서 학습하는 방법

Now

Future (?)

이미지 출처 : <https://www.datakeem.co/en/what-is-artificial-intelligence/>

머신러닝의 학습방법

머신러닝(Machine Learning)

10

▶ 머신러닝이란?

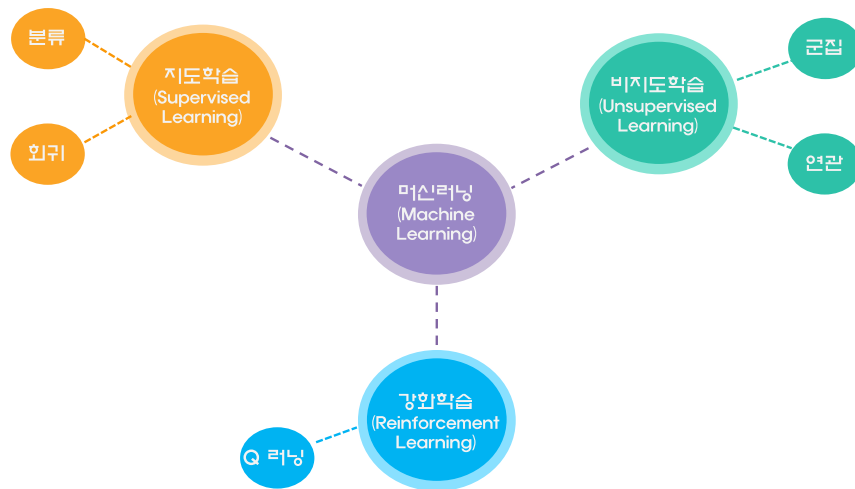
- ▶ 대량의 데이터와 알고리즘을 기반으로 경험을 통해 학습하여 데이터의 특징과 패턴을 발견
- ▶ 프로그래머가 명시적으로 코딩하지 않고 자기 개선과 예제를 통해 학습
- ▶ 학습으로 얻어진 정보를 기반으로 미래의 임의의 데이터에 대한 결과(값, 분포)를 예측

▶ 머신러닝에 필요한 주요 요소

- ▶ 다양한 형식의 데이터 (동영상, 이미지, 텍스트, 로그, html, xml, xlsx 등)
- ▶ 알고리즘 : 컴퓨터에게 지능(생각)을 심어주는 부분
 - ▶ 수학적 통계 모델, 의사결정트리, 서포트벡터머신, 신경망 등

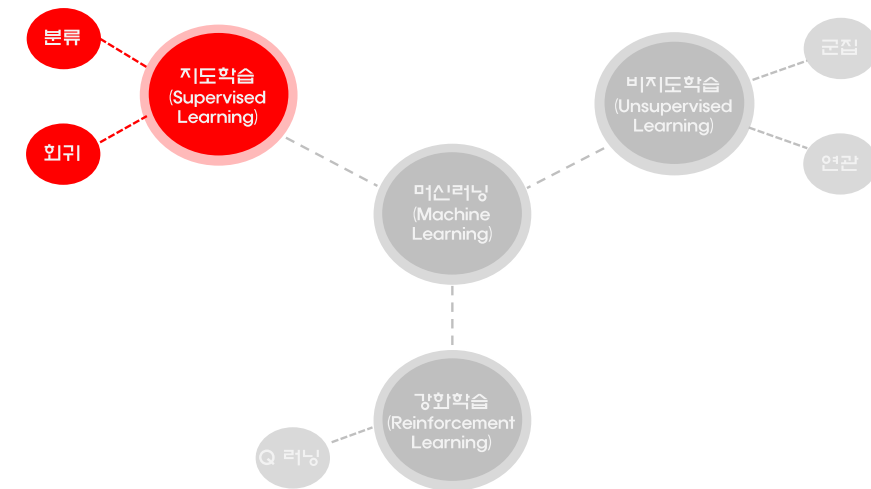
머신러닝의 학습 방법

11



머신러닝의 학습 방법

12



지도학습

13

▶ 지도학습 (Supervised Learning)

- ▶ 학습할 데이터와 명시적인 정답(레이블)을 이용해 데이터의 특성과 분포를 학습하여 새로운 데이터에 대해 미래결과를 예측하는 방법
- ▶ 과거의 데이터로부터 학습하여 결과를 예측하는데 주로 사용

▶ 분류(Classification)

- ▶ 학습데이터를 이용해 주어진 입력값이 어떤 종류인지 구별하는 것

▶ 회귀(Regression)

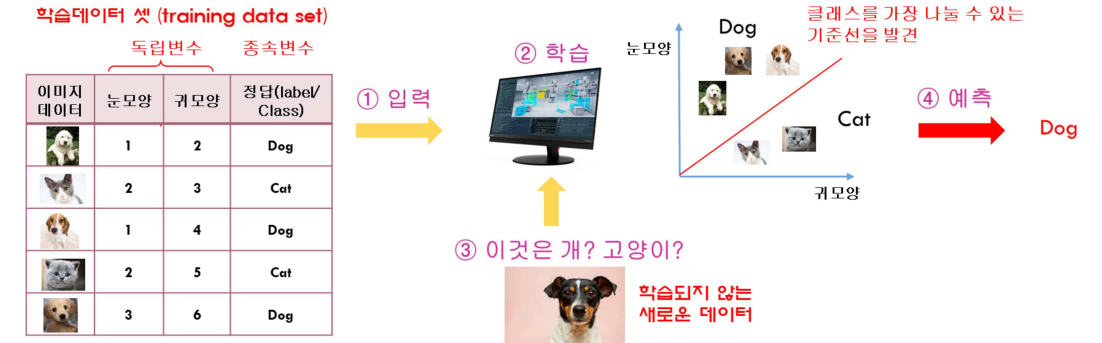
- ▶ 학습 데이터를 이용해 연속적인(숫자) 값을 예측하는 것

지도학습

14

▶ 분류(Classification)

- ▶ 개와 고양이 분류 예

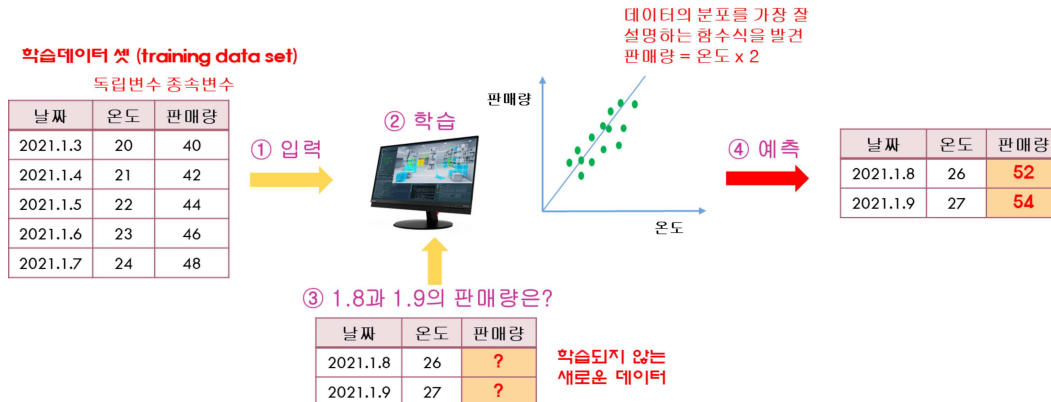


지도학습

15

▶ 회귀(Regression)

- ▶ 일별 온도와 레모네이드 판매량 예측 예



지도학습의 예

16

▶ 분류(Classification)

- ▶ 가지고 있는 데이터에 독립/종속 변수가 있고 종속변수가 이력(범주형 데이터)일 때 사용

독립변수	종속변수
공부시간	합격여부 (합격/불합격)
엑스선 사진과 영상 속 종양의 크기, 두께	악성 종양 여부(양성/음성)
폴중, 산도, 당도, 지역, 연도	와인의 등급
키, 몸무게, 시력, 지병	현역, 공역, 면제
메일발신인, 제목, 본문내용(사용 단어, 이모티콘 등)	스팸 메일 여부
고기의 지방함량, 지방색, 성숙도, 육색	소고기 등급

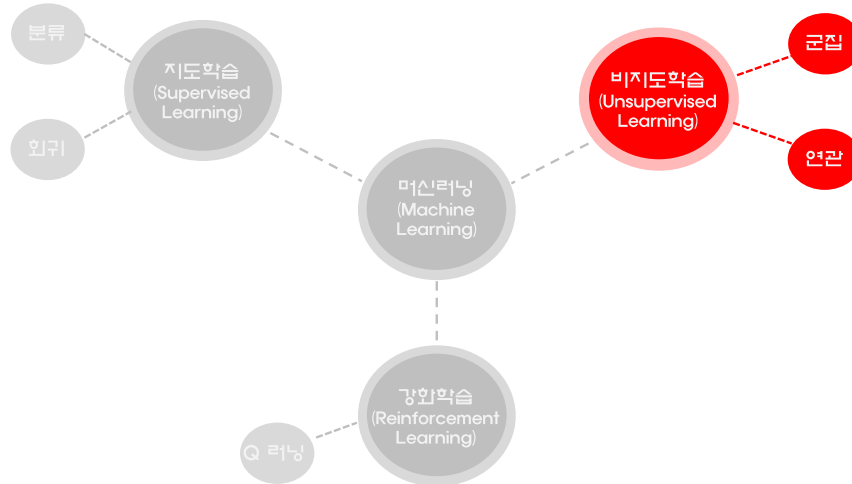
▶ 회귀(Regression)

- ▶ 가지고 있는 데이터에 독립/종속 변수가 있고 종속변수가 숫자(양적 데이터)일 때 사용

독립변수	종속변수
공부시간	시험점수
온도	레모네이드 판매량
역세권, 조망 등	집값
온실 기체량	기온 변화량
자동차 속도	충돌 시 사망 확률
나이	키

머신러닝의 학습 방법

17



비지도학습

18

비지도학습(Unsupervised Learning)

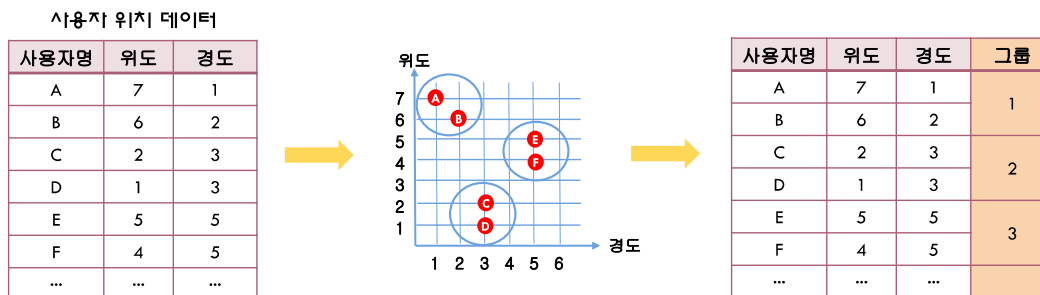
- ▶ 학습할 데이터에 **정답이 없이** 명확한 해를 계산하는 대신 **입력 데이터만으로 숨겨진 특징이나 패턴을 학습을 통해 발견하는 방법**
- ▶ **군집(Clustering)**
 - ▶ 사전 정보가 없는 주어진 데이터들에서 공통적인 요소들을 찾아 유사성이 높은 군들끼리 다른 그룹으로 나누는 방법
- ▶ **연관규칙(Association rule)**
 - ▶ 주어진 데이터에서 어떤 항목이 어떤 항목을 동반하여 등장하는지, 항목들 간에 연관규칙을 찾는 방법

비지도학습

19

▶ 군집(Clustering)

- ▶ 전국적으로 분포하는 사용자에게 최적의 배달을 위한 배달 본부 배치 예



비지도학습

20

▶ 연관규칙(Association rule)

- ▶ 음악추천 예

음악 구매 내역 (구매한 가수 1, 구매하지 않은 가수 0)

	사용자 1	사용자 2	사용자 3	사용자 4
BTS	1	0	0	1
브루노마스	1	0	0	0
아이유	1	0	0	1
마룬5	0	1	1	0
퀸	0	1	0	0

마룬5와 다른 가수의 유사도 계산

	BTS	브루노마스	아이유	마룬5	퀸
마룬5	0	0	0	1	0.71

‘마룬 5’ 를 구매한 사용자 3에게 ‘퀸’ 의 노래 추천

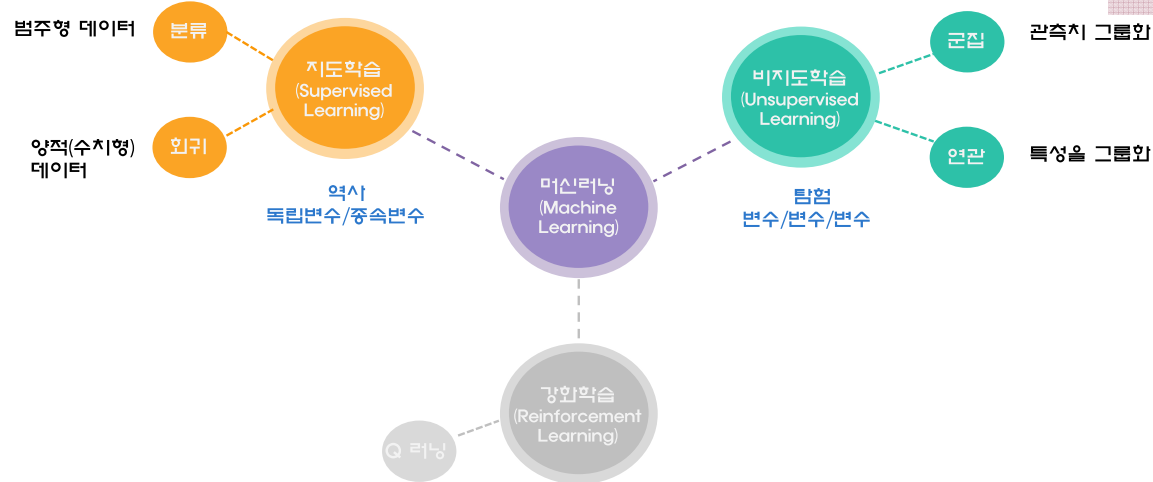
※ 마룬5와 다른 가수의 유사도 계산

$$\text{유사도} = \cos \theta = \frac{A \cdot B}{\|A\| \cdot \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

$$\text{코사인유사도} = \frac{0 \times 0 + 1 \times 1 + 1 \times 0 + 0 \times 0}{\sqrt{0^2 + 1^2 + 1^2 + 0^2} \times \sqrt{0^2 + 1^2 + 0^2 + 0^2}} = \frac{1}{\sqrt{2} \times \sqrt{1}} = 0.71$$

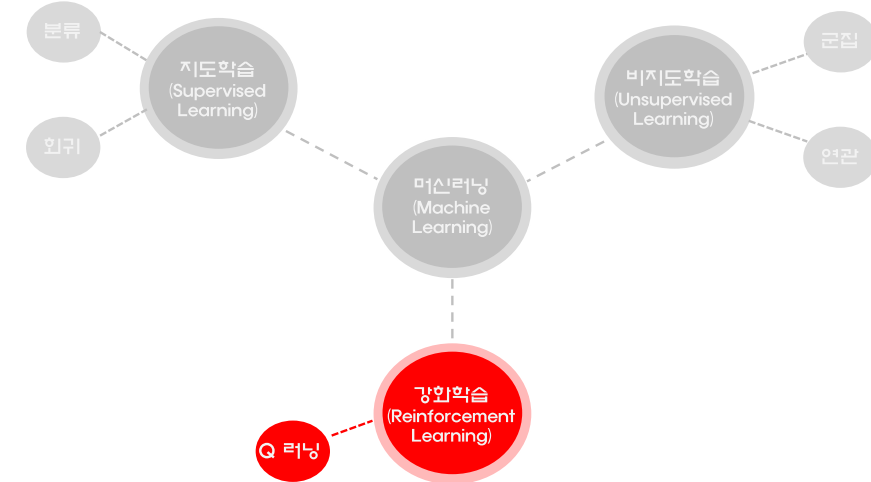
머신러닝의 학습 방법

21



머신러닝의 학습 방법

22



강화학습

23

▶ 강화학습 (Reinforcement learning)

- ▶ 학습의 주체가 어떤 환경에서의 시행착오의 경험을 통해 학습
- ▶ 지도학습과의 차이는 사람으로부터 학습을 받는 것이 아니라, 변화되는 환경으로부터 보상과 벌을 반복하며 학습한다는 차이를 가짐
 - ▶ 알파고 - 바둑의 기본 규칙과 자체 경기를 통해 습득한 3,000만 개의 기보를 학습한 후 스스로 대국하며 훈련하는 강화학습 알고리즘을 사용해 개발됨
 - ▶ 자율 주행 자동차와 드론 분야 등에서 활용

▶ Q-러닝

- ▶ 주어진 환경(state)에서 주체(agent)가 현재의 상태(state)를 관찰하며 선택할 수 있는 행동(action)들 중 가장 최대의 보상(reward)을 가져다주는 행동이 무엇인지 정책 데이터를 지속적으로 업데이트하는 알고리즘

Q 러닝

24

▶ 로봇에게 강화학습을 시키는 예

- ▶ 목표 : 출발점(S)에서 시작하여 도착점(G)까지 구멍(H)에 빠지지 않고 도착

현재상태 (State) : 16개의 격자 중 하나

학습주체 (Agent) : 로봇

행동 (Action) : 상하좌우 이동, 제자리

보상 (Reward) : 도착점에 도달 시 +1점



S	F	F	F
F	H	F	H
F	F	F	H
H	F	F	G

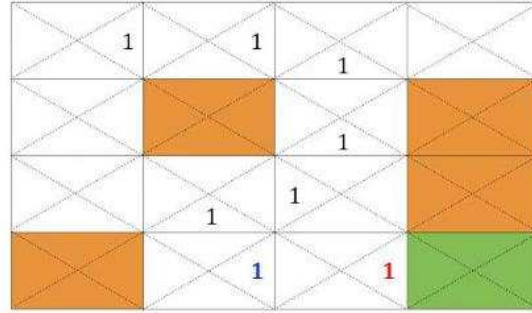
Q 러닝

25

▶ 로봇에게 강화학습을 시키는 예

▶ 보상과 벌점의 반복을 통해 경로를 학습

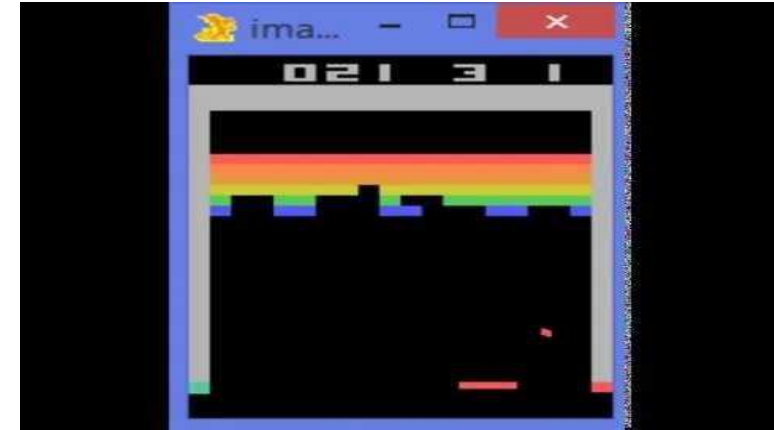
- 녹색박스 바로 앞에서 녹색격자로 이동한 마지막 행위에 +1
- 빨간 +1점의 격자로 들어서는 경우 +1점



강화학습 예

26

▶ 벽돌깨기 게임의 강화학습 예



<https://www.youtube.com/watch?v=V1eYniJQRnk>

다음시간에 배울 내용

27

주	주제	온라인	오프라인
1	인공지능의 과거 현재와 미래	1. 강의 및 교과목 소개(공통, 핵심만) 2. 인공지능의 과거와 현재 3. 인공지능의 미래와 다양한 시선 4. 인공지능 개발환경 구축과 사용법(Anaconda/Colab)	1. 강의 및 교과목 소개(분반별, 자세히) 2. 다양한 인공지능 기술 경험하기 (자연어처리, 시각, 음성) 3. 인공지능 챗봇만들기(IBM 왓슨 어시스턴트)
2	공공데이터를 이용한 사회문제 발견과 해결책 모색	1. 빅데이터의 정의와 가치 2. 공공데이터 수집하기 3. 공공데이터로부터 새로운 인사이트 발견하기 - 행정구역별 인구 데이터와 공공의료기관 현황 데이터 분석	1. 서울시 CCTV설치 현황 분석하기 2. 서울시 범죄발생 현황 분석하기
3	인공지능의 개요 및 머신러닝을 이용한 예측	1. 인공지능의 정의와 분류 2. 인공지능 학습방법 이해하기 3. 인공지능 알고리즘 소개	1. 머신러닝을 이용한 이미지 식별(구글 티쳐블 머신) 2. 머신러닝을 이용한 보스톤 집값 예측
4	인공지능과 데이터 윤리	1. 데이터의 불완전성과 결함에 따른 예측 오류와 차별 2. 데이터 왜곡에 따른 분석과 예측 결과 비교	1. 데이터 편향성이 예측에 미치는 영향 (구글 티쳐블 머신) 2. 데이터 왜곡에 따른 예측 결과 비교 - 타이타닉호 생존자 예측
5	인공지능과 알고리즘 윤리	1. 알고리즘 기반 의사결정 시스템의 한계 2. 윤리가 적용된 인공지능 알고리즘	1. 알고리즘에 따른 예측 결과 비교 - 보스톤 집값 예측 - 폐암환자 생존 여부 예측
6	인공지능에 대한 다양한 이슈와 우리의 자세 고찰	1. 인공지능의 윤리적/법적 쟁점 (자율주행자동차, AI로봇, 트랜스 휴먼 등) 2. 인공지능시대 사회, 경제적 불평등 문제 3. 인공지능과 프라이버시 4. 인공지능의 윤리적 대응과 규제	1. 자율주행 자동차의 행동학습 시나리오 경험하기 2. 비윤리적 데이터 생성과 수집(웹 크롤링을 이용한 데이터 수집)
7		기말고사	