

01 인공지능이란?

목차

1. 머신러닝의 종류
2. 강화학습과 신경망
3. 이 책의 구성

1. 머신러닝의 종류

- 인공지능(Artificial Intelligence)이란 인공적으로 만들어진 지능
- 인공지능에는 전문가 시스템(Expert System), 데이터 마이닝(Data Mining), 패턴인식(Pattern Recognition), 로봇틱스(Robotics) 등과 같이 여러 분야가 있으며, 머신러닝(Machine Learning)도 그 중 하나

- **머신러닝의 종류**

- 머신러닝은 주어진 데이터를 이용해 목적에 맞게 예측하거나 새로운 규칙을 찾도록 학습 대상을 학습시키는 것
- 학습 방법에 따라 그림 1.1과 같이 지도학습(Supervised Learning), 비지도학습(Unsupervised Learning), 강화학습(Reinforcement Learning)으로 구분

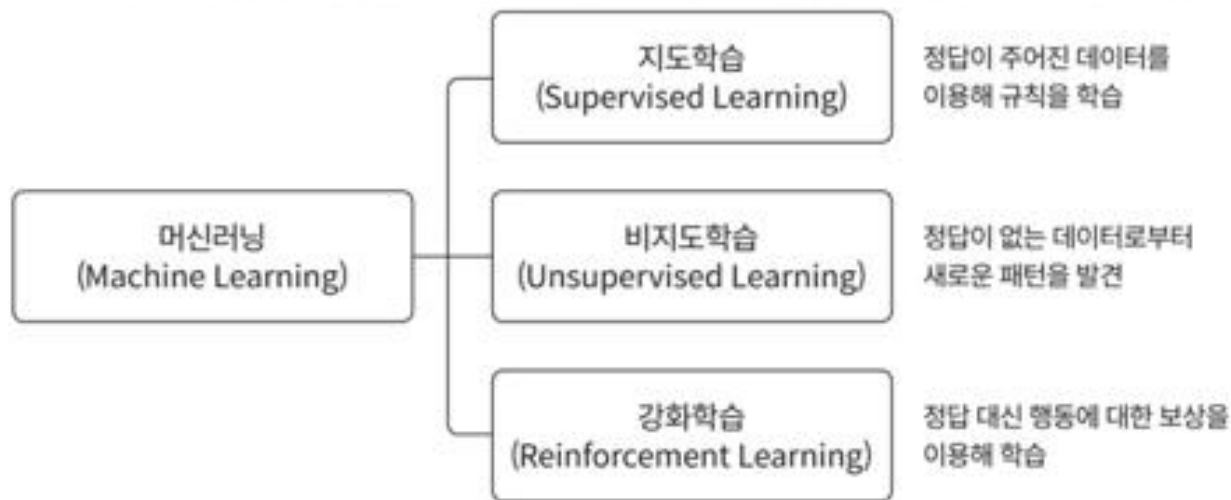


그림 1.1 머신러닝의 종류

1. 머신러닝의 종류

• 지도학습

- 원하는 방향으로 가르치고 이끄는 지도(指導)
- 선생님이 정답을 가르친다고 해서 교사학습이라고도 불림
- 문제와 문제에 대한 정답이 쌍으로 이뤄진 학습 데이터를 사전에 준비. 문제와 함께 정답을 줘서 문제에 대해 정확하게 정답을 출력할 수 있도록 반복해서 학습. 학습이 끝나면 학습에 사용되지 않은 **새로운** 데이터에 대해 예측, 분류 등을 수행
- 학습 대상이 학습하기 위해 주는 문제와 정답 데이터 쌍을 훈련 데이터(Training Data), 학습 후 성능을 테스트하기 위해 주는 데이터를 테스트 데이터(Test Data)
- 대표적인 지도학습 알고리즘으로 의사결정나무(Decision Tree), 신경망(Neural Network), 선형회귀(Linar Regression) 등이 존재

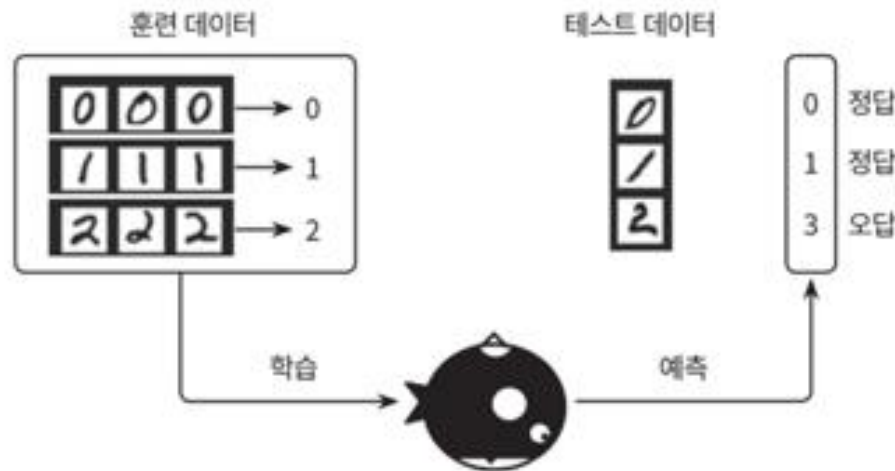


그림 1.2 지도학습

1. 머신러닝의 종류

• 비지도학습

- 비지도학습은 데이터로부터 일정한 규칙을 찾아내는 학습 방법
- 정답이 주어지지 않기 때문에 데이터에 따라 여러 가지 새로운 규칙 획득 가능
- 비지도학습은 정답이 없는 데이터로부터 규칙을 찾는데 사용되며, 대표적인 알고리즘으로 주성분 분석(Principal Component Analysis), K-평균(K-Means), 자기조직화지도(Self-Organizing Map)이 있음

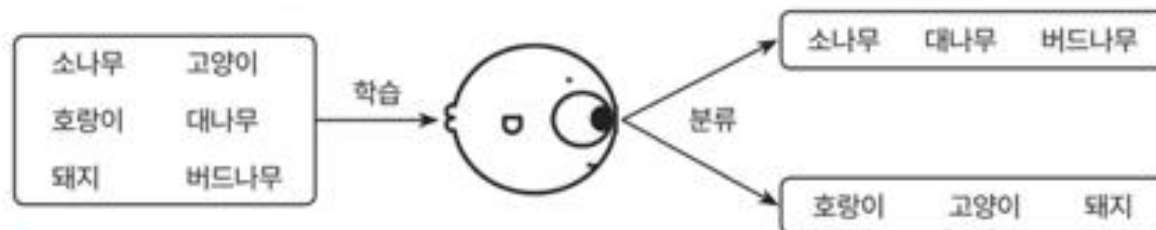


그림 1.3 비지도학습

1. 머신러닝의 종류

• 강화학습

- 강화학습에는 보상(Reward)과 벌칙(Punishment)을 정의
- 어떤 상태에서 취한 행동이 잘된 행동이면 보상을 주고, 잘못된 행동이라면 벌칙을 줌
- 학습 대상은 시행착오(trial-and-error)를 통해 스스로 벌칙을 최소화하고 보상을 최대화하도록 행동을 학습. 학습 대상이 이용할 수 있는 정보는 보상 뿐이기 때문에, 학습 목적에 맞는 보상을 설정하는 것이 매우 중요
- 강화학습은 직접적으로 정답을 주지 않지만 주어진 보상을 최대화할 수 있도록 시행착오를 통해 행동을 최적화

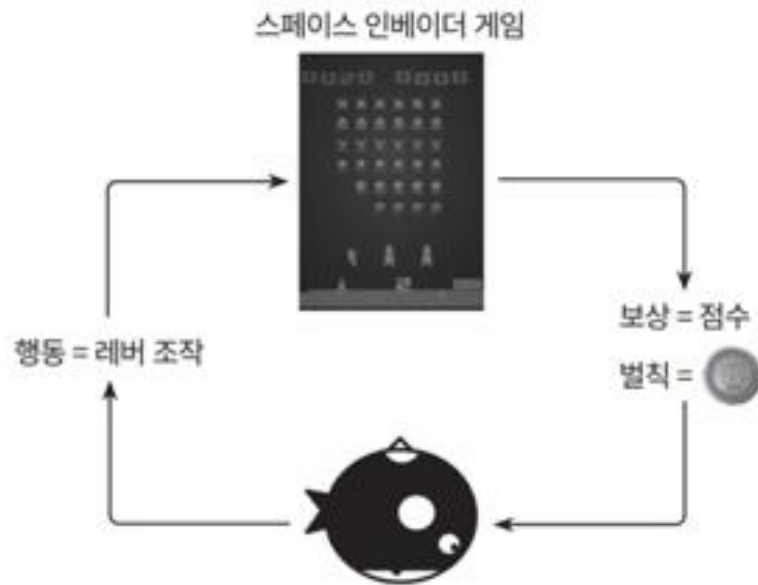


그림 1.4 강화학습¹

2. 강화학습과 신경망

- **강화**란 어떤 지속적인 자극에 의한 반응이 점차적으로 강해지거나 약해지는 것을 의미
- 초기에는 간단한 산업용 로봇의 제어, 휴머노이드 로봇의 행동학습, 자율자동차의 경로 학습 등 다양한 분야에서 사용되기 시작
- 강화학습의 대표적인 알고리즘은 Q-learning
- 강화학습의 영역이 확대되면서 휴머노이드 로봇의 행동 학습, 자율 주행 자동차와 같이 상태공간이 큰 문제에 적용되면서 지금까지 문제되지 않았던 메모리의 한계에 부딪히기 시작. 강화학습 연구자들은 이런 문제를 극복하기 위해 메모리를 이용하는 Q-table 대신 함수근사(Function Approximation)에 눈을 돌리기 시작
- 강화학습을 연구하던 연구자들은 강화학습의 학습 방법 최적화와 함께 어떻게 함수를 근사를 할 것인가를 동시에 연구해야 하는 상황에 처함

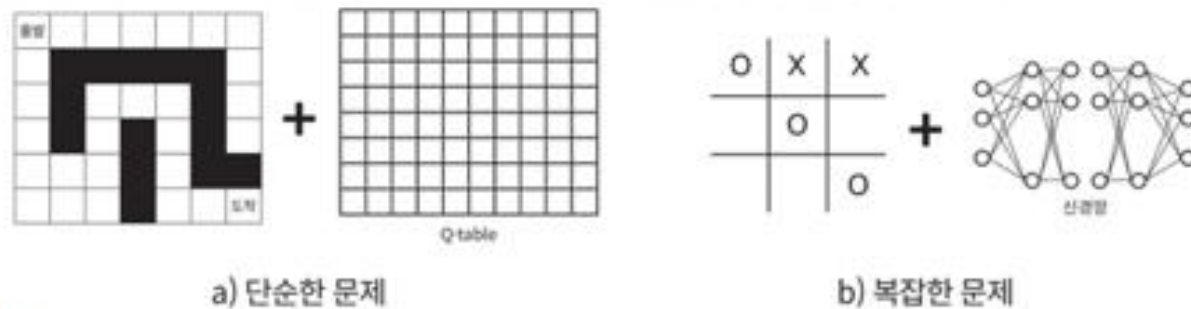


그림 1.5 Q-table과 신경망 이용

2. 강화학습과 신경망

- 오차 역전파법(Back-propagation)의 등장과 함께 신경망(Neural Network)이 제2의 부흥기를 맞게 되고, 신경망으로 다차원의 비선형 함수를 근사할 수 있다는 연구가 알려지면서 강화학습 연구자들 사이에는 그림 1.5의 b)와 같이 Q-table을 신경망으로 대체하는 연구가 활발하게 진행
- RBF(Radial Basis Function) 네트워크나 LVQ(Learning Vector Quantization) 네트워크 같은 다양한 형태의 신경망도 이 당시에 활발하게 사용. 그러나 풀고자 하는 문제가 커질수록 신경망이 복잡하고 깊어지면서 신경망의 약점인 경사소실(vanishing gradient) 문제가 생기기 시작하고, 한계가 드러나면서 신경망의 인기는 점점 떨어짐. Q-table의 대체 수단으로 신경망을 사용하던 연구자들은 신경망의 한계로 다른 수단을 찾기 시작
- **2016년 3월** 한국의 바둑기사 이세돌 9단과 인공지능이 대국한다는 소식은 인공지능 연구자들뿐 아니라 전 세계 사람이 주목

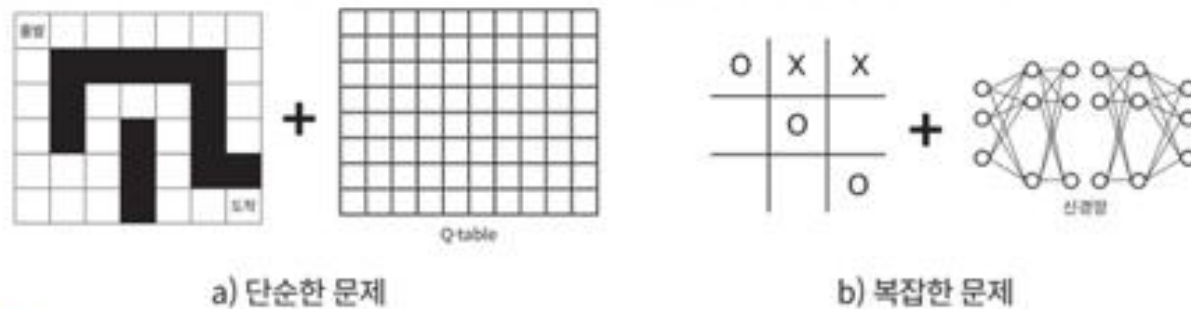


그림 1.5 Q-table과 신경망 이용

2. 강화학습과 신경망

- 2016년 3월 한국의 바둑기사 이세돌 9단과 인공지능이 대국한다는 소식은 인공지능 연구자들뿐 아니라 전 세계 사람이 주목
- 알파고의 학습에 강화학습을 사용했다는 것이 알려지면서 다시 강화학습이 주목받기 시작. 물론 알파고가 강화학습만을 사용한 것은 아님. 신경망의 약점이었던 경사소실 문제를 제프리 힌튼(Geoffrey Hinton)이 "A fast learning algorithm for deep belief nets"라는 논문으로 해결 가능하다는 것을 소개한 것을 시작으로 다시 신경망의 부흥기가 시작
- 2012년 국제이미지인식기술대회(ILSVRC) 대회에서 제프리 힌튼 팀이 알렉스넷(AlexNet)을 이용해 뛰어난 성능으로 우승하며 신경망의 부활 알림

2. 강화학습과 신경망

- 기존의 뉴런으로만 연결된 신경망에서 이미지 처리를 추가한 합성곱 신경망 (Convolutional Neural Netowrks)이라는 이름으로 이미지 인식과 분류 분야에서 독보적인 존재로 자리 잡음
- 부활한 신경망을 강화학습에서 다시 사용할 수 있게 되고, 복잡도가 약 10^{171} 에 달해서 절대 정복할 수 없다는 바둑의 상태공간을 결과적으로 그림 1.6과 같이 신경망을 이용해 구현 가능

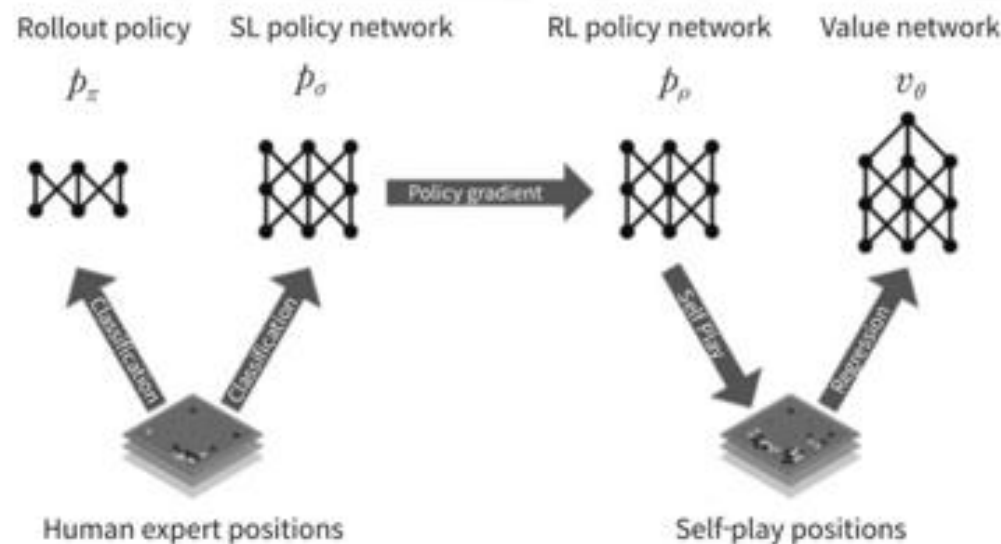


그림 1.6 알파고의 신경망 학습²

3. 이 책의 구성

- 2장

- 강화학습의 기본 개념부터 여러 가지 가오하학습 알고리즘을 에이전트의 미로 탐색 예제를 이용해 파이썬 코드와 함께 설명

- 4장

- 틱택토(Tic Tac Toe) 게임을 플레이하는 강화학습 기반의 인공지능 플레이어를 구현
- 틱택토는 복잡도 측면에서는 바둑보다 덜하지만 인공지능 플레이어를 만드는데 괜찮은 주제
- 사람과 대전하는 인공지능 플레이어로 랜덤 플레이어, 몬테카를로 플레이어, Q-learning 플레이어 순으로 구현

- 이 책에서 사용된 예제코드는 파이썬 3.6.8, 텐서플로 1.6.0

- 다음 사이트에서 코드를 내려 받을 수 있음

- <https://github.com/wikibook/rlnn>

- 이제 강화학습부터 시작해보자.