심층 강화 학습 기술 동향

□ 김중헌 / 고려대학교

요약

강화 학습 기술은 많은 분야에서 매우 적극적으로 활용되는 기계 학습 기술 중의 하나이며 최근 이를 사용한 많은 연구 결과를 다양한 기관에서 활발하게 보여주고 있다. 본 고에서는 이러한 강화 학습 기술에 대한 기본적인 소개와 해당 기술의 심층 강화 학습으로의 발전에 대해서 논한다. 더불어 이러한 심층 강화 학습의 많은 분야 중에서 최근 활발히 논의되는 모방 학습에 대해서 알아보고 그 활용성에 대해서 논한다.

I. 서 론

최근 다양한 분야에서 강화 학습 기술을 사용하는 데에 많은 관심이 모아지고 있다. 네트워크 관리, 로봇 제어, 지능형 게임 등 다양한 분야에서 강화 학습은 활발하게 활용되고 있다. 위와 같은 강화 학습 기술은 일반적인 딥러닝 알고리즘과 같이 한 번의 의사 결정으로 그 결과가 나오는 것이 아니라 연속적인 의사 결정이 있는 경우에 주로 사용이 된다.

이러한 강화 학습은 다양한 응용 분야에서 활발히 사용되고 있으며 2016년 알파고 이후 많은 발전을 이루어오고 있다. 본 고에서는 이러한 강화 학습 알고리즘이 기계 학습이라는 연구 분야 내에서 어떠한 위치를 가지고 있는지와 다른 기계 학습 알고리즘 대비 어떠한 점이 다르고 어떠한 점이 특징인지를 논한다. 그 후에 고전적인 강화 학습 알고리즘의 개념과 그 한계에 대해서 서술한다. 이러한 한계를 극복하기 위해서 강화 학습 알고리즘은 딥러닝/인공 신경망을 통하여 행동을 도출하는 방식인 심층 강화 학습으로 발전한다. 해당 심층 강화 학습의 구체적인 동작방식과 한계에 대해서 논하며, 최근 심층 강화 학습 연구에서 급격한 발전을 이루어오고 있는 모방 학습에 대해서 알아보고 그 응용에 대해서 본 고에서는 알아본다.

II. 기계 학습 분류

기계 학습은 인간이 학습하는 모습을 모방하여 이

를 통계학 혹은 수학적으로 표현하여 기계 혹은 컴퓨터가 인간의 학습하는 과정을 계산적으로 모사하는 방법론을 연구하는 학문이다. 본 기계 학습 연구 영역은 “지도 학습(Supervised Learning)”, “비지도 학습(Unsupervised Learning)”, 그리고 “강화 학습(Reinforcement Learning)”이라는 세 가지 카테고리로 나눌 수 있다. 각각의 항목에 대해서는 세부적으로 다음에서 알아보도록 한다.

1. 지도 학습

지도 학습(Supervised Learning)은 학습을 수행할 수 있는 데이터와 그 데이터의 이름표에 해당하는 레이블이 주어진 상태에서 학습을 수행하는 방식을 의미한다. 예를 들어 사과와 바나나를 구별하는 분류기 인공지능 알고리즘을 설계한다고 하자. 이 때에 사과의 모양을 학습시키기 위한 사진들과 바나나의 모양을 학습시키기 위한 사진들을 다수 입력하여 그들의 통계적인 일반화를 추론하게 유도한다. 예를 들어, 비록 단 하나의 사과도 정확한 빨간색이면서 완벽한 구의 모습을 하고 있지 않더라도, 빨갛고 동그란 모양이면 사과라 인식하고, 유사하게 단 하나의 바나나도 정확한 노란색이면서 타원형으로 길쭉한 모습을 하고 있지 않더라도, 노란색 길쭉한 모양이면 바나나로 인식하게 된다. 이러한 학습의 과정에서 사과의 사진을 보여주면서 사과라고 알려주는 레이블이 있어야 사과의 사진이 사과 모습임을 인식하게 된다. 반대로 사과 사진을 보여주며 바나나라고 칭한다면 거꾸로 학습이 될 것이다.

즉 지도 학습이란 데이터와 그 데이터를 설명하는 레이블이 다수 존재할 때에 이 둘의 요소를 가지고 학습을 진행함을 의미한다. 지도 학습의 가장 대표적인 예로써 위에서 전술한 바와 같이 분류기 설계이다.

2. 비지도 학습

비지도 학습(Unsupervised Learning)은 지도 학습의 반대말이므로 데이터가 없는 상태에서 학습을 진행한다고 생각할 수 있다. 그러나 비지도 학습은 데이터가 없는 상태에서 학습을 수행하는 것이 아니라 데이터는 주어지지만 그 데이터를 설명하기 위한 레이블/이름표는 존재하지 않는 것을 의미한다.

가장 대표적인 예로 아이에게 펜을 줄 때에 아이가 이를 먹으려 입에 대는지 아닌지에 대한 의사 판단이 비지도 학습에 해당한다. 아이에게 펜을 줄 때에 “이것은 먹을 수 없으므로 먹지 말 것”을 강조하더라도 아이는 말을 알아들을 수 없기 때문에 이러한 정보를 숙지할 수 없다. 즉 레이블을 부여하지 않은 것과 동일한 효과를 가진다. 이러한 상황에서 아이는 기본적인 본능만 존재하는 상태이므로 식욕에 근거하여 펜을 입에 가져가 댈 것이다. 이를 통하여 입에 가져다 댄 물체가 그동안 섭취한 음식들과 비슷하면 그 펜을 음식으로 인식하고 먹으려 할 것이다. 이렇게 펜을 입에 가져다 대는 행위는 펜을 1차원 벡터로 표현한 상태에서 기존에 섭취했던 음식물들의 벡터들과 유사도 판단을 하는 것이다. 만약에 유사도가 높다고 나온다면, 현재 입에 대는 물체가 기존에 섭취했던 음식물들과 비슷함을 의미하고 따라서 먹으려 할 것이다.

이러한 과정은 1) 물에의 벡터로 표현, 2) 기존의 벡터들과 유사도 판단을 통한 특정 군집에 포함되는지의 여부 판단으로 요약할 수 있으며, 이는 추천 시스템이나 검색 엔진 설계에서 가장 기본적으로 사용되는 방식이다. 추천 시스템에서는 새로운 아이템이 시스템에 추가되었을 때에 사용자에게 해당 아이템을 추천할지 말지는 그동안 사용자가 좋아하던 아이템들과의 유사도 판단을 통해 결정할 수 있고, 검색 엔진 설계에서는 검색 대상이 되는 웹 문서들에서 키워드를 추출한 후에 이를 벡터화하고, 주

어진 쿼리 역시 벡터화 한 후에 쿼리의 벡터와 웹 문서의 벡터들 간의 유사도 판단을 수행한 후에 유사도가 높은 순서대로 출력해 준다. 즉 추천 시스템 설계와 검색 엔진 설계는 비지도 학습의 가장 대표적인 예시이다.

3. 강화 학습

강화 학습(Reinforcement Learning)은 기계 학습의 한 종류로 어떠한 환경에서 어떠한 행동을 했을 때 그것이 잘된 행동인지 잘못된 행동인지를 나중에 판단하고 보상을 함으로써 반복을 통해 스스로 학습하게 하는 분야이다[1]. 강화 학습에는 에이전트(Agent)와 환경(Environment)이라는 두 가지 구성 요소가 존재한다.

에이전트는 특정 환경에서 자신의 행동(Action)을 결정하고 환경은 그 결정에 대한 보상을 부여한다. 이 보상은 행동 즉시 결정되기보다는 여러 행동들을 취한 후에 한꺼번에 결정된다. 이는 특정 행동을 취했을 때 바로 그 행동에 대한 평가를 내릴 수 없는 경우가 많기 때문이다.

이와 같은 과정을 통하여 강화 학습이 기존의 지도 및 비지도 학습과 근본적으로 다른 점은 한 번의 행동으로 결과가 나오는 것이 아니라 연속적인 행동을 도출하는 것이라는 것이다. 이러한 특성에 근거하여 다양한 시변환 환경에 적응하며 매번 해당 환경에 맞는 의사 결정을 해야 하는 통신 시스템이나, 상대방이 수를 두는 것에 따라 적응하며 대응하는 게임 알고리즘, 그리고 로봇 제어 등에 주로 활용된다.

III. 딥러닝 연산 과정

1. 개요

딥러닝은 인공 신경망이라 불리우는 구조를 근간으로 사용한다. 이러한 인공 신경망은 분류 등의 여러 통계적인 기능을 수행할 때에 은닉 계층(Hidden Layer)을 사용함으로써 비선형성을 증가시킨다. 즉 은닉 계층이 없다면 일반적인 선형 분류기/인식기와 동일하다. 더불어 이러한 은닉 계층이 인공 신경망에서 매우 많이 쌓여 있다면 딥러닝이라 칭한다.

즉 딥러닝은 인공 신경망 구조를 기본으로 하며 이는 지도 학습의 범주에 속한다. 왜냐하면 인공 신경망을 학습시킨다고 함은 인공 신경망의 입력층과 출력층에 각각 데이터와 레이블을 입력하기 때문에 지도 학습의 학습 원리와 동일하다.

그러나 딥러닝에는 지도 학습의 요소만 존재하는 것은 아니다. 적대 생성망(Generative Adversarial Network (GAN))이라 불리우는 기술은 기존에 없는 데이터를 새로 생성하는 것을 그 목적으로 하며 이러한 경우에는 비지도 학습의 모습을 일부 차용한다. 더불어 전술한 바와 같이 강화 학습도 인공 신경망으로 연산하는 방식도 있기 때문에 딥러닝은 강화 학습의 모습 역시 일부 차용한다.

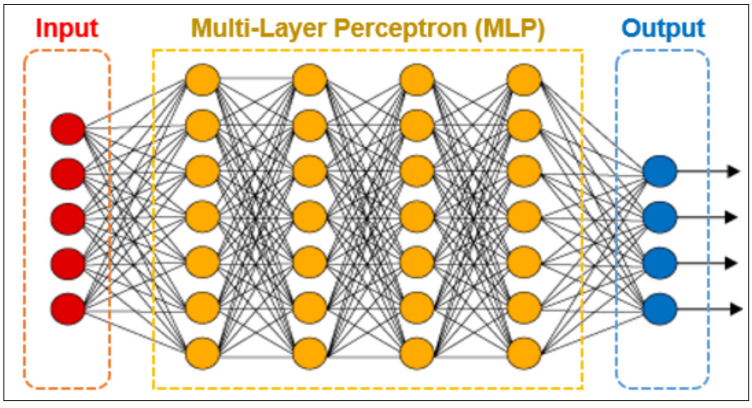
결과적으로 인공 신경망을 근간으로 하는 딥러닝은 지도 학습을 기본으로 하여 비지도 학습과 강화 학습의 일부를 차용함으로써 기계 학습 알고리즘 중에서 최근 가장 많은 주목을 받고 있다.

2. 연산 과정

딥러닝이라 불리우는 많은 은닉 계층을 가지는 인공 신경망 구조는 다음의 세 단계를 통하여 학습과 추론을 수행한다.

1단계) 모델 구성: 주어진 문제에 가장 적합한 인공 신경망 구조를 정의하는 과정이다.

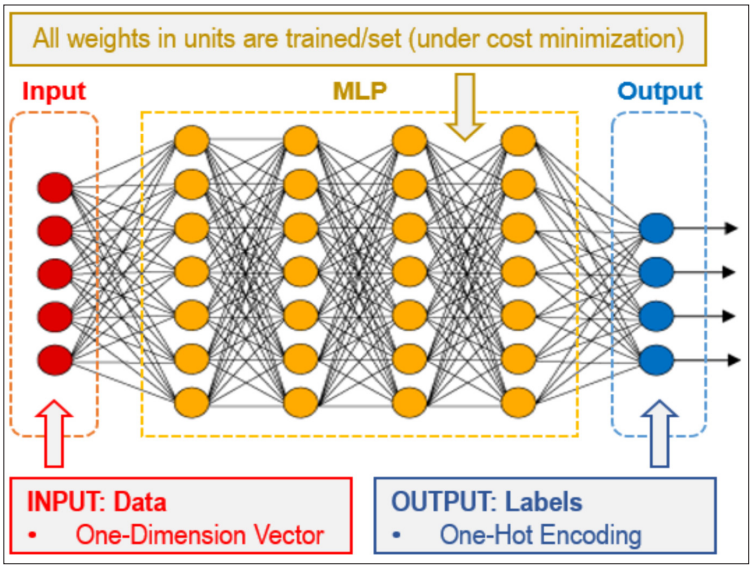
주어진 문제가 매우 복잡하여 많은 연산을 요구하



<그림 1>

면 은닉 계층을 늘림과 동시에 은닉 계층에서의 유닛의 개수도 늘린다. 〈그림 1〉에서는 노란색 부분이 은닉 계층을 의미하며 세로로 4줄이 존재한다 함은 은닉 계층의 개수가 총 4개임을 의미한다. 그리고 각 은닉 계층은 7개의 유닛이 존재한다. 만약에 비선형성을 늘려 성능을 높이고 싶다면 이 은닉 계층의 수를 늘리고 좀 더 정교한 연산을 요구한다면 유닛의 개수를 증가시킨다. 그리고 〈그림 1〉은 가장 기본적인 인공신경망 구조이며, 시각 정보 처리를 위한 딥러닝 구조를 설계한다면 Convolutional Neural Network (CNN)라는 딥러닝 구조를 사용하여 2차원의 이미지나 3차원 비디오 정보를 처리함에 효율을 높인다. 만약에 시계열 정보를 사용한다면 시간 정보를 포함하여 학습을 수행 가능한 Recurrent Neural Network (RNN)라는 딥러닝 구조를 사용하여 학습을 진행한다. 이렇듯 주어진 문제에 맞는 인공 신경망 딥러닝 구조로 모델을 구성한다.

2단계) 모델 학습: 위의 1단계에서 구성한 모델을 많



<그림 2>

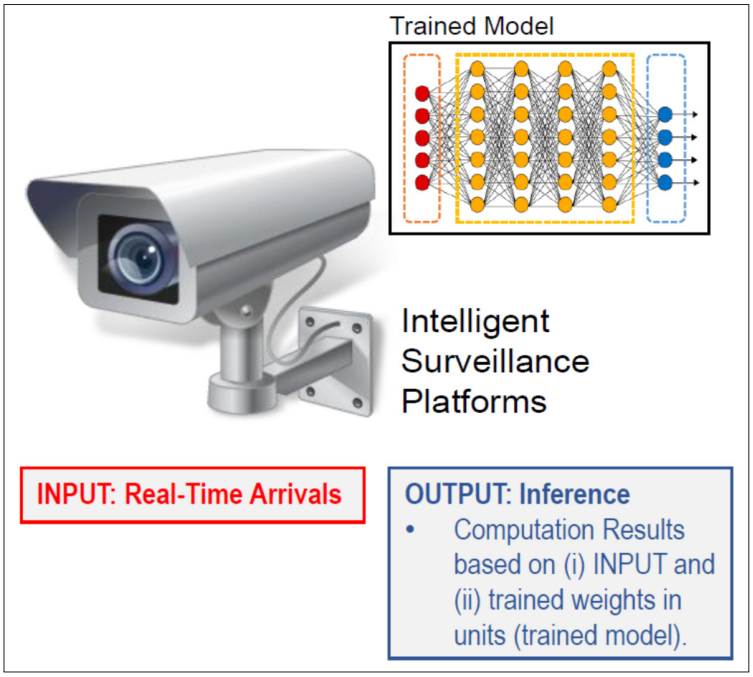
은 양의 데이터와 그의 레이블을 가지고 학습을 진행한다.

1단계에서 구성한 모델에 입력과 출력 많은 양의 데이터를 기반으로 하여 학습을 진행한다. 이와 같이 학습을 진행하면 중간에 은닉 계층에 있는 파라미터 값이 세팅이 되면서 학습이 진행된다. 만약에 우리가 구성한 인공 신경망 모델이 사과와 바나나를 구별하는 식별자라면 입력은 사과나 바나나의 이미지가 될 것이고, 출력은 2차원 벡터가 되어 [0,1]이면 사과를 의미하는 이름표/레이블이고 [1,0]이면 바나나를 의미하는 이름표/레이블이다.

3단계) 추론 : 많은 양의 데이터를 기반으로 하여 은닉 계층의 파라미터가 학습이 된 결과를 바탕으로 하여 실제 데이터가 들어올 때에 결과를 추론한다.

〈그림 3〉에서 보는 바와 같이 실제 학습된 모델을 가지고 실생활에 적용할 때에 실제 사과/바나나 식별자라면, 사과 혹은 바나나의 사진을 보여주면 학습된 은닉 계층의 파라미터를 기반으로 확률적으로 사과인지 바나나인지를 도출한다. 예를 들어 이 값이 [0.3, 0.7]이면 바나나일 확률이 70%임을 의미하므로 바나나라고 추론하게 된다.

위와 같은 세 단계가 딥러닝 연산에 기본적인 학습과 추론 과정에 해당한다. 수학적으로 볼 때에 은닉 계층의 수가 많고 개별 은닉 계층에서의 유닛의 수가 많을수록 좀 더 복잡하고 정교한 분류가 가능하다. 딥러닝은 이러한 은닉 계층이 깊게(Deep) 20여개 정도 쌓은 것을 말하며 이러한 다수의 은닉 계층에 근거하여 매우 정교하고 우수한 성능을 보인다.

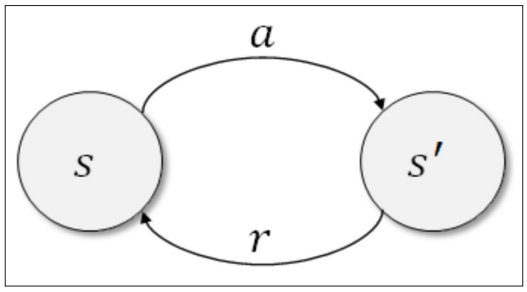


<그림 3>

IV. 강화 학습 기본과 한계

1. 기본 개념

이전 II.3장에서 알아본 바와 같이 강화 학습은 한 번의 행동으로 끝나는 기계 학습 연산이 아니라 행동의 연속이 있는 기계 학습 연산이다. 구체적으로 각각의 행동에 따른 보상이 존재할 때에, 연속적인 행동에 따른 보상의 합의 평균이 최대가 되는 행동을 찾는 과정이다.



위의 그림에서 현재 강화 학습을 동작시키는 에이전트(Agent)가 S라는 공간에 있다고 하자. 이 에이전트는 다양한 행동(Action)이 가능하며 그 중에서 위의 그림의 a라는 행동을 선택하면 r만큼의 보상(Reward)을 받는다. 그 후에 s′라는 다음 공간으로 이동한다. 그렇다면 이러한 관계는 다음과 같은 식으로 표현 가능하다.

$$Q(s,a) = r + \max_{a'} Q(s',a')$$

위의 식에서 좌변에 해당하는 Q(s,a)는 S라는 공간에서 a라는 행동을 통해 받는 보상을 의미한다. 즉 Q라는 함수는 s와 a라는 값에 의해 바뀔 수 있는 보상을 의미한다. s라는 공간에서 a라는 행동을 취했으므로 r이라는 보상을 부여받으므로 우변에 r이 있으며, 그 후에 s′라는 새로운 공간으로 이동한 후에 그 후에도 반복적으로 최대의 보상을 주는 행동을 반복적으로 수행할 것이다. 따라서 우변은 위와 같이 표현된다. 위의 식에서 보듯이 강화 학습은 반복적인 연산을 기반으로 하고 있기 때문에 수학적으로 점화식에 근거한다. 즉 위의 식에서 보듯이 좌변의 Q값을 구하기 위해서 우변에도 Q가 존재하기 때문에 점화식에 근거하여 연산된다.

위와 같은 기본적인 강화 학습에 이론적인 근거를 두고 개발된 알고리즘은 Q-Learning 및 마코프 의사 결정 과정(Markov Decision Process) 등이 있다[2].

2. 한계

위의 강화 학습의 근간을 이루는 점화식을 푸는 가장 대표적인 알고리즘은 동적 계획법(Dynamic Programming)이다. 동적 계획법의 연산 복잡도는 Pseudo-Polynomial로써 이는 문제의 상태의 수가 적을 때에는 최적의 해를 주어진 시간 내에 연산할 수 있으나 그렇지 못할 경우에는 매우 많은 연산이 필요함을 의미한다. 고전적인 강화 학습 문제는 상태의 수가 크지 않으므로 일반적인 동적 계획법으로 연산 결과를 도출함에 큰 문제가 없으며 최적의 해도 도출할 수 있다. 그러나 최근 강화 학습 알고리즘이 사용되는 예시들은 알파고 등 매우 많은 상태 정보를 가지고 있는 문제들이다. 따라서 일반적인 동적 계획법으로 답을 도출하는 것은 주어진 시간 내에 불가능하다는 것이 근본적인 한계이다.

V. 심층 강화 학습 개요

위의 장에서 알아본 바와 같이 Q-Learning과 마코프 의사 결정 과정 같은 고전적인 강화 학습 알고리즘들은

점화식에 근거하고 있으며 이는 동적 계획법에 의해 연산된다. 동적 계획법의 Pseudo-Polynomial 복잡도에 근거하여 상태의 수가 적을 때에는 최적의 해를 도출하지만, 상태의 수가 커지게 되면 연산이 불가능하다. 따라서 최근 알파고와 같은 복잡한 게임 알고리즘이나 복잡한 통신 시스템에 자동 제어 같은 문제에는 고전적인 강화 학습의 사용이 불가능하다.

따라서 딥러닝 인공 신경망 구조를 차용하여 인공 신경망의 입력과 출력에 각각 상태 정보와 그에 따른 행동 정보를 넣어 학습한다. 그러면 새로운 상태가 입력되었을 때에 학습에 근거하여 새로운 행동 정보를 도출한다.

예를 들어, 이세돌 선수와 같이 바둑을 두는 심층 강화 학습 엔진을 만든다고 하자. 인공 신경망을 중간에 두고 입력 부분에는 이세돌 선수가 지금까지 대국을 둔 바둑판 정보를 입력하고, 출력 부분에는 해당 각각의 바둑판에서 대국을 둔 수를 입력한다. 학습이 끝나고 해당 인공 신경망이 새로운 바둑판을 입력으로 받으면 학습에 근거하여 이세돌 선수의 행동을 근사(Approximation)하여 행동을 도출한다.

이와 같이 인공 신경망 기반의 강화 학습을 딥러닝(Deep Learning)으로 강화 학습(Reinforcement Learning)을 수행한다고 하여 심층 강화 학습(Deep Reinforcement Learning)이라고 한다[3]. 본 알고리즘은 딥러닝 학습 과정의 2단계의 학습 단계에서는 시간이 많이 걸릴 수 있더라도 실제 해당 심층 강화 학습 모델이 사용되는 3단계의 추론에서는 인공 신경망 기본 연산만 하면 되므로 고전적인 강화 학습보다 실제 상황에서는 더욱 빠르게 결론을 도출한다. 그러나 고전적인 강화 학습이 상태의 수가 적다면 최적의 해를 도출할 수 있는 반면, 심층 강화 학습은 인공 신경망 학습이 충분히 고성능으로 이루어지지 않는다면 성능의 열화를 가져올 수 있다. 따라서 학습의 성능이 매우 중요하다고 할 수 있다.

VI. 모방 학습 및 응용

1. 기본 개념

모방 학습은 이전 장에서 다룬 심층 강화 학습의 연구 분야 중에서 최근 많이 다루어지는 분야이다. 본 모방 학습은 다양한 상황(혹은 상태)에 따른 전문가의 행동을 데이터 집합 D로 구성한다. 이러한 집합을 인공 신경망 학습을 위한 데이터로 하여 인공 신경망을 학습하여 활용하는 것이 모방 학습이다[4, 5].

2. 응용

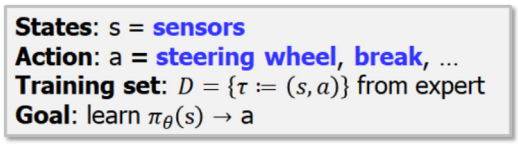
모방 학습이 대표적으로 사용된 분야이자, DeepMind가 모방 학습을 활용하여 실증한 가장 대표적인 예시는 게임이다[6]. DeepMind는 스타크래프트II의 게임을 할 수 있는 딥러닝 시스템을 구축함에 있어서 모방 학습을 활용하는 방식으로 하였다.

States: s = **minimap**, **screen**
Action: a = **select**, **drag**
Training set: $D = \{\tau := (s, a)\}$ from expert
Goal: learn $\pi_\theta(s) \rightarrow a$

위의 그림에서 보듯이 상태(State)정보로 스타크래프트 게임의 미니맵(minimap)과 스크린(screen)을 입력받고, 행동(Action)으로써 프로게이머들의 유닛 선택(select)과 이동(drag)을 입력한다. 프로게이머로부터 수집한 이러한 상태와 행동을 학습 데이터 집합(Training set)으로 만들어 강화 학습 에이전트를 학습시키고, 새로운 상태인 s를 입력받을 때에 이러한 학습에 근거하여 행동 a를 도출한다.

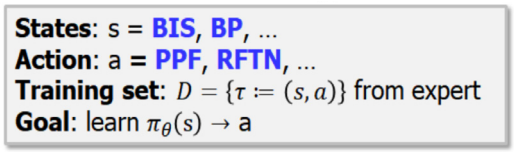
더불어 가장 많이 언급되는 응용 분야는 자율 주행

이다[4, 5].



위의 그림에서 보듯이 상태(State)정보는 차량의 LIDAR 및 mmWave 등 센서들로 입력받은 정보들이다. 행동(Action)으로써 위와 같은 상태 정보에 맞는 운전대 회전(steering wheel)과 브레이크를 밟는 강도(break)를 입력한다. 해당 차량의 차주의 운전으로 인하여 수집한 이러한 상태와 행동을 학습 데이터 집합(Training set)으로 만들어 강화 학습 에이전트를 학습시키고, 새로운 상태인 s를 입력받을 때에 이러한 학습에 근거하여 행동 a를 도출한다.

다음으로 마취 의료 전문가 시스템이다[7].



위의 그림에서 보듯이 상태(State) 정보는 환자의 심박수(BIS)와 혈압(Blood Pressure)으로 입력받은 정보들이다. 행동(Action)으로써 위와 같은 상태 정보에 맞는 마취과 의사의 프로포폴(PPF)과 레미펜타닐(RFTN)이라는 마취약의 투약의 적정량 정보를 입력한다. 해당 의료 행위로 인하여 수집한 이러한 상태와 행동을 학습 데이터 집합(Training set)으로 만들어 강화 학습 에이전트를 학습시키고, 새로운 상태인 s를 입력받을 때에 이러한 학습에 근거하여 행동 a를 도출한다. 그러나 이와 같은 시스템은 행동을 취하자마자 상태 정보가 바로 변화하지 않으므로 지연이 존재하는 Delayed 시스템이므로 그에 대한 고려가 함께 필요하다.

3. 특징

위와 같은 모방 학습은 특정 상황에서 나오는 행동이 전문가로써 특이한 경우에 학습의 결과도 전문가의 행동처럼 나와야 의미가 있다. 동일한 상태에서 다양한 행동이 가능하게 학습 데이터 집합이 구성된다면 위와 같은 모방 학습이 의미가 없어진다. 따라서 모방 학습은 전문가 시스템의 설계에서 가장 큰 의미가 있다.

VII. 결 론

본 고에서는 기계 학습의 분류와 강화 학습에 대해서 고찰하고, 고전적인 강화 학습이 가지는 한계를 극복하기 위하여 딥러닝 기반의 강화 학습인 심층 강화 학습 알고리즘의 기본 개념에 대해서 논한다. 더불어 최근 많은 발전을 이루어 온 모방 학습 기술에 대해서 논하고 그에 따른 응용 분야를 고찰한다.

참 고 문 헌

[1] M. Shin, D.-H. Choi, and J. Kim, "Cooperative Management for PV/ESS-Enabled Electric-Vehicle Charging Stations: A Multiagent Deep Reinforcement Learning Approach," IEEE Transactions on Industrial Informatics, vol. 16, no. 5, May 2020.

[2] M. Choi, A. No, M. Ji, and J. Kim, "Markov Decision Policies for Dynamic Video Delivery in Wireless Caching Networks," IEEE Transactions on Wireless Communications, vol. 18, no. 12, December 2019.

[3] S. Jung, W.J. Yun, M. Shin, J. Kim, and J.-H. Kim, "Orchestrated Scheduling and Multi-Agent Deep Reinforcement Learning for Cloud-Assisted Multi-UAV Charging Systems," IEEE Transactions on Vehicular Technology, vol. 70, no. 6, June 2021.

[4] M. Shin and J. Kim, "Randomized Adversarial Imitation Learning for Autonomous Driving," in Proc. International Joint Conference on Artificial Intelligence (IJCAI), Macau, China, August 2019.

[5] W.J. Yun, M. Shin, S. Jung, S. Kwon, and J. Kim, "Parallelized and Randomized Adversarial Imitation Learning for Safety-Critical Self-Driving Vehicles," Journal of Communications and Networks, vol. 24, no. 3, June 2022.

[6] W.J. Yun, S. Yi, and J. Kim, "Multi-Agent Deep Reinforcement Learning using Attentive Graph Neural Architectures for Real-Time Strategy Games," in Proc. IEEE International Conference on Systems, Man, and Cybernetics (SMC), Melbourne, Australia, October 2021.

[7] M. Shin and J. Kim, "Joint Behavioral Cloning and Reinforcement Learning Method for Propofol and Remifentanil Infusion in Anesthesia," in Proc. IEEE International Conference on Information Networking (ICOIN), Jeju, Korea, January 2021.

필 자 소 개

김 중 헌

- 2004년 : 고려대학교 컴퓨터학과 학사
- 2006년 : 고려대학교 컴퓨터학과 석사
- 2006년 ~ 2009년 : LG전자 멀티미디어연구소 주임연구원
- 2014년 : University of Southern California, Computer Science 박사
- 2013년 ~ 2016년 : 미국 인텔 본사연구소 Systems Engineer
- 2016년 ~ 2019년 : 중앙대학교 소프트웨어대학 조교수
- 2019년 ~ 현재 : 고려대학교 전기전자공학부 부교수
- 주관심분야 : 딥러닝, 심층 강화 학습, 강화 학습 기반 시스템 제어