



## 2024-1학기 창의학기제 최종결과보고서

창의 과제	국문	강화학습의 이론적 탐구 및 알고리즘의 실질적 구현을 통한 기법별 실제 성능 비교분석			
	영문	Theoretical exploration of reinforcement learning and practical implementation of algorithms to compare the performance of different techniques in practice			
학습기간	2024.03.04 ~ 2024.06.16				
유형(■)	<input type="checkbox"/> 개발형 <input type="checkbox"/> 창업형 <input type="checkbox"/> 창작형 <input type="checkbox"/> 문제해결형 <input checked="" type="checkbox"/> 탐구형				
※ 수강학점에 따른 총 학습시간 및 학습회차 10회 이상 준수					

### 1. 학습 목표

- DQN, DDQN, Dueling DQN과 같은 주요 강화학습 알고리즘들에 대한 심층적인 이해
- DQN, DDQN, Dueling DQN 등의 알고리즘들을 구현 및 성능 비교분석
- 실제 환경에서 강화학습 알고리즘을 어떻게 확장할 수 있을지에 대한 탐구
- 각 알고리즘의 성능을 향상하는 방법에 관해 탐구

### 2. 학습 내용

- DQN, DDQN, Dueling DQN에 대한 개념을 이해하고 OpenAI GYM에서 제공하는 강화학습 환경을 사용하여 직접 코드를 구현해 보았다. 최종적으로 성능 비교를 통해 이론적으로 이해한 내용을 통계적으로, 시각적으로 확인함으로써 알고리즘에 대해 심화된 이해를 할 수 있었다. 먼저 DQN, DDQN, 그리고 Dueling DQN에 대해 기본적인 개념을 이해하고 그 알고리즘들이 나온 논문을 읽었다. 이 논문들을 통해 심층적으로 알고리즘들을 이해할 수 있었고 논문에서 나온 실험을 보고 추후에 알고리즘 간 성능 비교를 어떻게 해야할지 대략적으로 생각해볼 수 있었다. 이후에는 간단한 Cartpole 예제를 통해 세 가지 알고리즘을 구현해 보았고, 성능 비교를 위해 더 복잡한 예제인 Lunar Lander 실험까지 진행하였다. Lunar Lander 게임 환경에서 DQN, DDQN, 그리고 Dueling DQN 알고리즘을 반복적으로 진행해서 실험 결과를 통계 내고 시각화하여 이론적으로 공부한 내용을 정량적으로, 정성적으로 확인하였다.

### 3. 회차별 학습보고

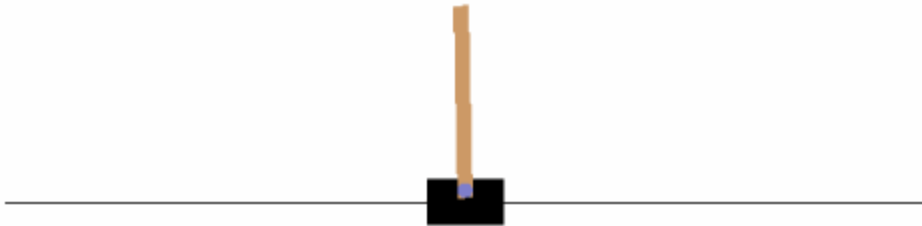
회차	학습 목표 및 활동	학습방법	학습시간
1	DQN, DDQN, Dueling DQN 관련 개념 숙지	논문 탐색 및 구글링	6
2	추가 개념 정리 및 팀원들과 개념 교류	세미나를 통한 개념 교류	6
3	선행 연구 파악	논문 정리 및 개념 교류	6



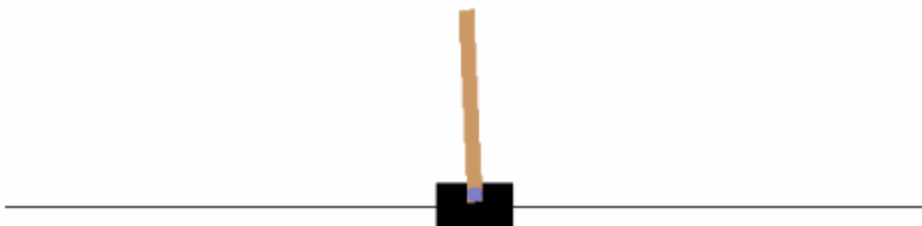
4	선행 연구 파악	논문 정리 및 개념 교류	6
5	카트폴 예제 환경 설정	인터넷 자료와 해당 코드 참고	6
6	카트폴 코드 구현 및 리뷰	인터넷 자료와 해당 코드 참고	6
7	모델(DQN, Doble DQN, Dueling DQN) 코드리뷰	인터넷 자료와 해당 코드 참고	6
8	실험 결과 통계 및 시각화	코드 실행 및 정리	6
9	카트폴 예제에 대한 한계점 탐구	결과 분석과 토의 진행	6
10	루나랜더 예제 환경 설정	인터넷 자료와 해당 코드 참고	6
11	루나랜더 코드 구현 및 리뷰	인터넷 자료와 해당 코드 참고	6
12	모델(DQN, Doble DQN, Dueling DQN) 코드리뷰	인터넷 자료와 해당 코드 참고	6
13	실험 결과 통계 및 시각화	코드 실행 및 정리	6
14	실험 결과 분석	결과 분석과 토의 진행	6
15	최종 결과물 정리	세미나를 통해 최종 정리	6

#### 4. 최종결과물 및 팀별 성과 \*구체적으로 명시

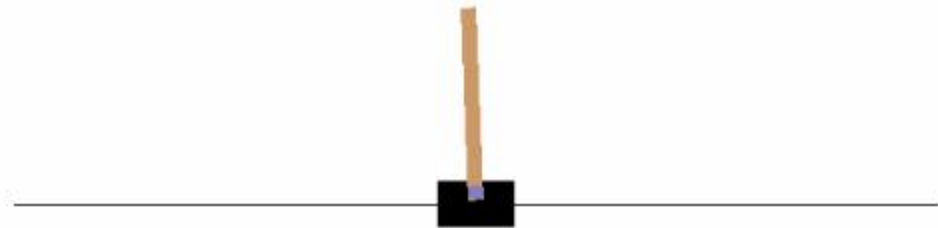
(※ 학습 성과를 증빙할 수 있는 그림, 사진 등 성과물)



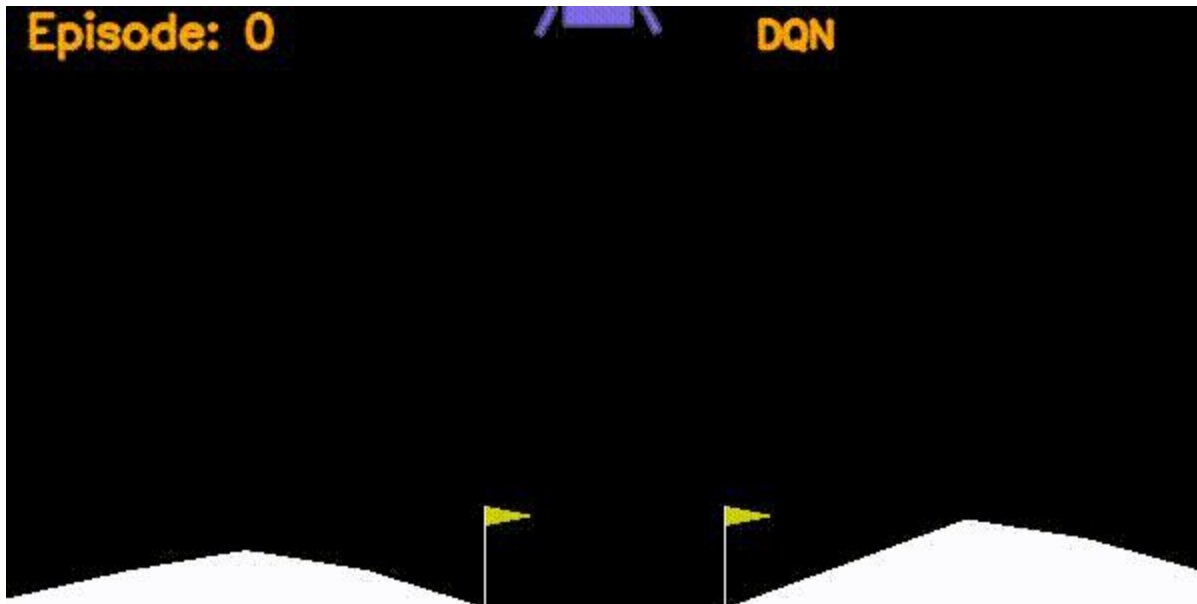
- DQN 알고리즘을 이용하여 학습한 모델이 Cartpole 게임에서 성공 조건을 10번 연속으로 성공한 것을 애니메이션으로 시각화



- DDQN 알고리즘을 이용하여 학습한 모델이 Cartpole 게임에서 성공 조건을 10번 연속으로 성공한 것을 애니메이션으로 시각화



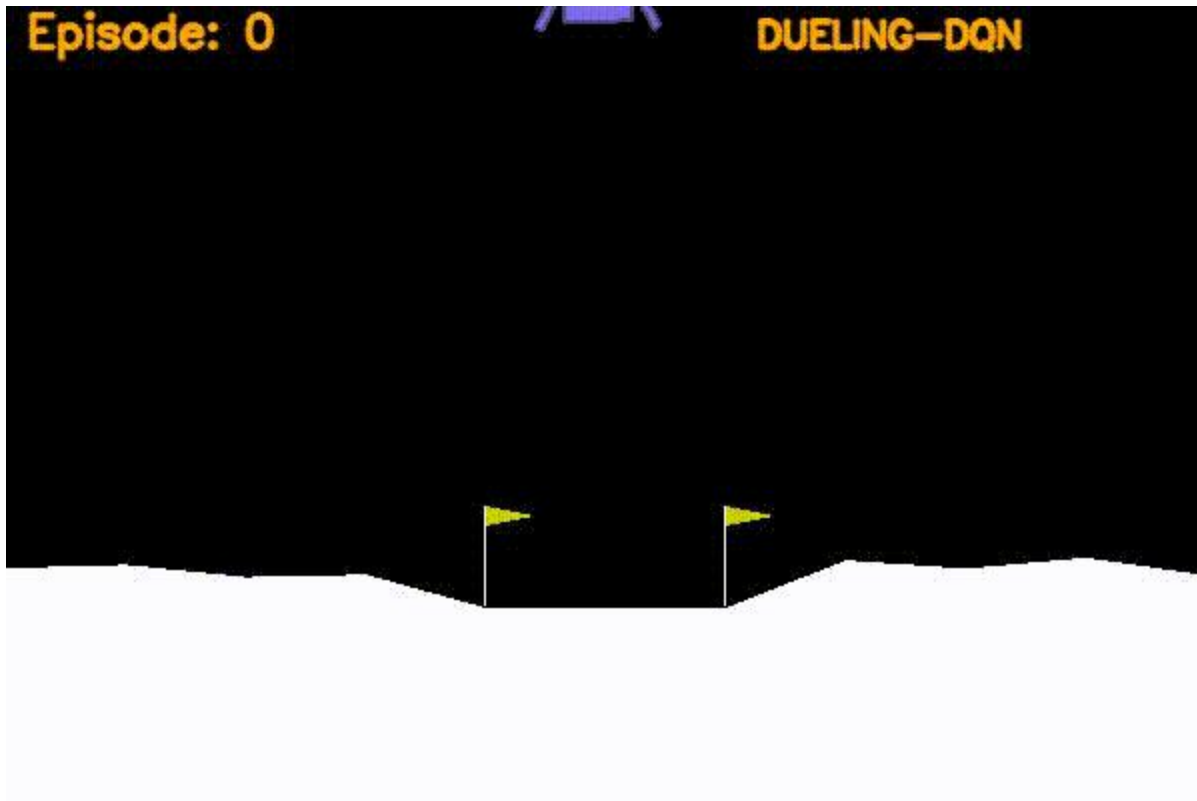
- Dueling DQN 알고리즘을 이용하여 학습한 모델이 Cartpole 게임에서 성공 조건을 10번 연속으로 성공한 것을 애니메이션으로 시각화



- DQN 알고리즘을 이용하여 학습한 모델을 2000 에피소드 동안 학습시킨 후 5번의 에피소드 동안 테스트한 것을 애니메이션으로 시각화



- DDQN 알고리즘을 이용하여 학습한 모델을 2000 에피소드 동안 학습시킨 후 5번의 에피소드 동안 테스트한 것을 애니메이션으로 시각화



- Dueling DQN 알고리즘을 이용하여 학습한 모델을 2000 에피소드 동안 학습시킨 후 5번의 에피소드 동안 테스트한 것을 애니메이션으로 시각화



DQN 알고리즘에 비해 DDQN 알고리즘은 실험을 여러 번 진행해도 Average Score가 큰 편차 없이 비슷하게 나타난 것을 확인했다. 이를 통해 특정 조건에서 Q-value의 overestimate이 일어나는 DQN 알고리즘을 보완한 DDQN의 강건성을 확인할 수 있었다. Dueling DQN은 DQN보다 훨씬 적은 적은 step으로 에피소드 종료 조건을 성공한 것을 확인했다. 이를 통해 advantage function과 value function을 나누어 구한 다음 합쳐서 Q-value를 구하는 방식이 효율적인 것을 확인하였다.

## 5. 기대효과

- ◆ 창의 학습을 통해 다양한 강화학습 알고리즘에 대한 개념적인 이해를 넘어서 직접 코드를 구현함으로써 심층적인 이해를 얻었다.
- ◆ 코드 구현을 통해 각 알고리즘의 작동 원리, 수렴 방식, 하이퍼파라미터의 영향 등을 체계적으로 이해하였고 강화학습 분야의 심층적인 이해가 가능하였다.
- ◆ 알고리즘을 구현하고 실험을 진행하는 과정에서 프로그래밍 및 실험 능력이 향상되었다. 또한, 실험 설계와 실행에 필요한 능력이 발전되었다.
- ◆ 알고리즘이 동작한 것을 수치로 확인하고 시각화하여 이해함으로써, 이론적인 내용을 통계적으로 분석하는 능력을 향상시켰다.
- ◆ 간단한 예제를 통해 알고리즘의 특성 및 응용 방안을 모색해보므로써 추후 실제 응용에서 어떤 방식으로 사용될지에 대해 고려할 수 있고, 해결책을 모색함으로써 이에 대한 대처 능력이 향상되었다.
- ◆ 팀원들과 교류하며 학습을 수행함으로써 소통 능력이 향상되고, 효율적으로 팀 프로젝트를 수행하는 법을 모색할 수 있었다.