

DeepRacer 개요

동양미래대

목차

1. AWS DeepRacer 둘러보기
2. Reinforcement Learning 기초

학습목표

- AWS DeepRacer 의 구성에 대해 이해한다.
- Reinforcement Learning 에 대한 기초적 이론 지식을 쌓는다.

1. AWS DeepRacer 둘러보기

1. AWS DeepRacer 둘러보기

1) 자율주행 차량 훈련하려면

- 자율주행을 할 모델을 훈련하려면 자율주행 시뮬레이터가 필요합니다.
(Why) 사고위험, 효율성
- 실제 세상과 비슷한 시뮬레이터 환경에서 자율주행 모델을 훈련시키고
훈련된 모델을 실제 차량에 장착하여 자율주행을 할 것입니다.

가상 공간에서 모델 훈련



실제 공간에서 자율주행



훈련된 모델을 오프라인 차량에 장착

1. AWS DeepRacer 둘러보기

1) 자율주행 차량 훈련하려면

처음부터 이 모든 걸 직접 한다면?

- 시뮬레이터 제작
- 강화학습 모델 구축
- 모델 훈련할 컴퓨터 자원 마련
- 오프라인 차량 제작
-



1. AWS DeepRacer 둘러보기

2) AWS DeepRacer

- AWS DeepRacer 는 자율주행을 구현하기 위한 모든 과정을 손 쉽게 할 수 있도록 구현한 서비스



- **AWS (Amazon Web Services) ?**
: 아마존닷컴의 클라우드 컴퓨팅 서비스
- **클라우드 컴퓨팅 서비스 ?**
: IT 인프라 구축에 필요한 온갖 서비스

[AWS DeepRacer 메인 화면]
<https://aws.amazon.com/ko/deepracer/>

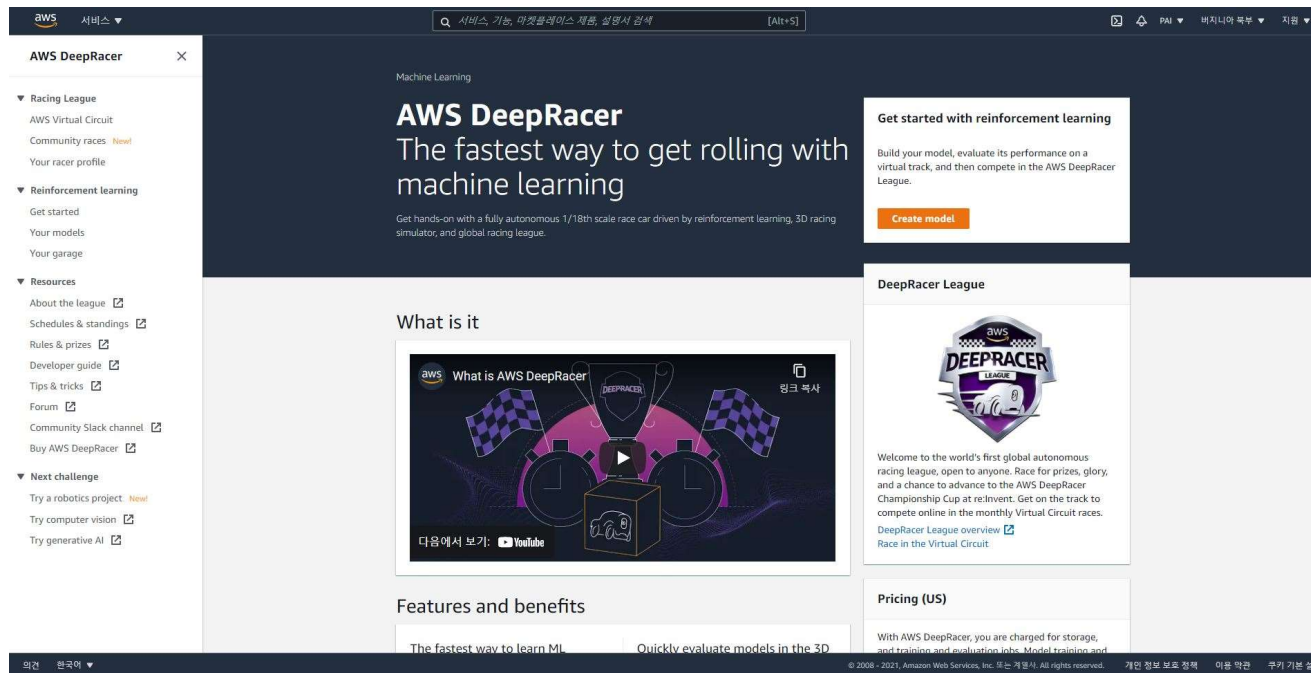


1. AWS DeepRacer 둘러보기

3) DeepRacer 콘솔 들어가기

(방법 1) AWS 홈페이지 -> 콘솔에 로그인 (우측 상단) -> 서비스 (좌측 상단) -> Machine Learning -> AWS DeepRacer

(방법 2) console.aws.amazon.com/deepracer



1. AWS DeepRacer 둘러보기

4) DeepRacer 콘솔 메뉴

▼ Racing League

AWS Virtual Circuit

전 세계 사람들과 가상 레이스

Community races **New!**

친구 및 지인들과 가상 레이스

Your racer profile

레이싱 프로필 관리

▼ Reinforcement learning

Get started

강화학습에 대해 배워보기

Your models

훈련된 모델 관리

Your garage

차량 에이전트 관리

▼ Resources

About the league [↗](#)

리그 소개

Schedules & standings [↗](#)

일정 및 순위

Rules & prizes [↗](#)

규칙 및 상품

Developer guide [↗](#)

개발자 가이드

Tips & tricks [↗](#)

팁

Forum [↗](#)

토론

Community Slack channel [↗](#)

Buy AWS DeepRacer [↗](#)

오프라인 차량 구매

▼ Next challenge

Try a robotics project **New!**

Try computer vision [↗](#)

Try generative AI [↗](#)

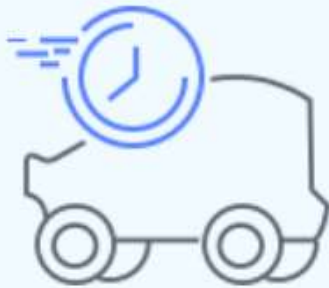
1. AWS DeepRacer 둘러보기

5) Race Type



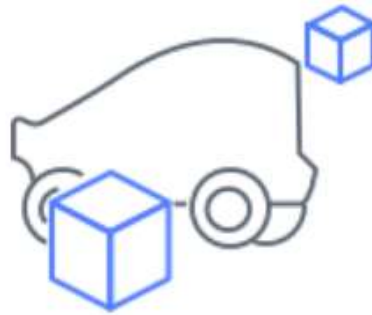
Time trial

The agent races against the clock on a well-marked track without stationary obstacles or moving competitors.



Object avoidance

The vehicle races on a two-lane track with a fixed number of stationary obstacles placed along the track.



Head-to-head racing

The vehicle races against other moving vehicles on a two-lane track.



1. AWS DeepRacer 둘러보기

6) Vehicle



DeepRacer

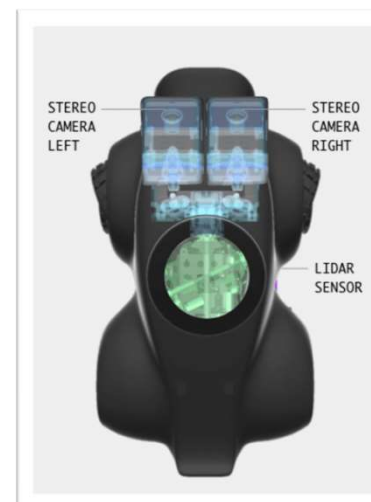
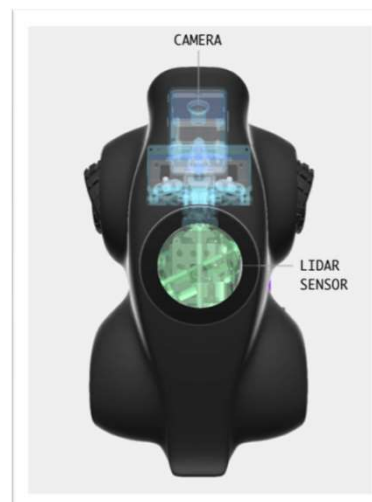
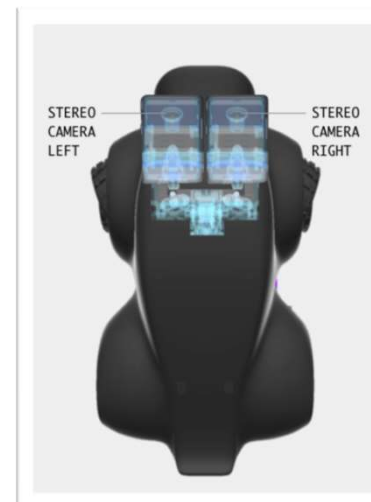
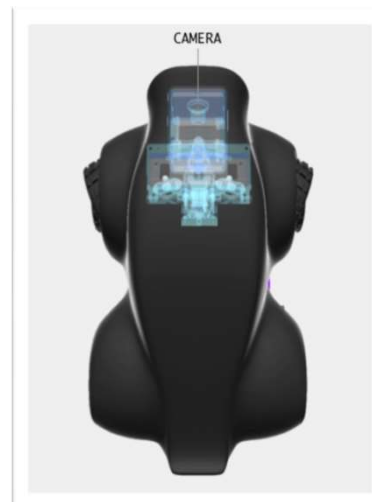
\$379.05

DeepRacer + Sensor Kit
(Evo)

\$568.10

Sensor Kit

\$236.55



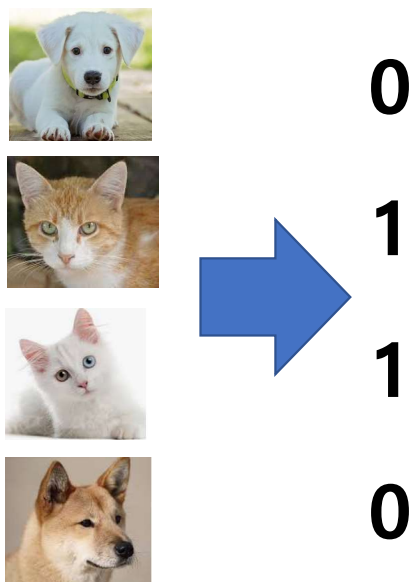
2. Reinforcement Learning

2. Reinforcement Learning

1) 머신러닝 종류

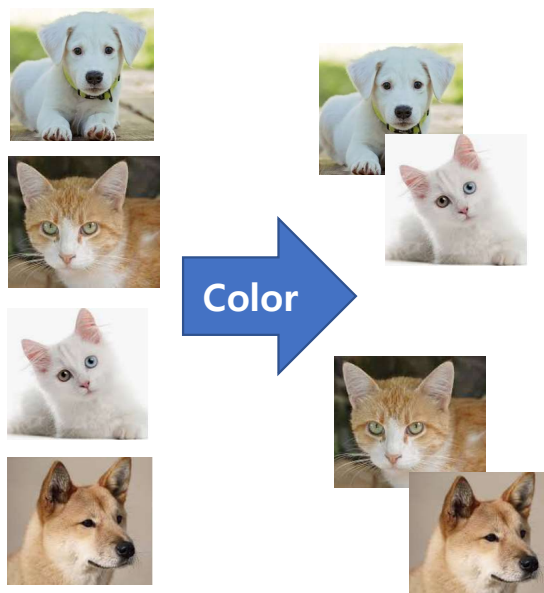
지도 학습 (Supervised learning)

- 입력 (X) 과 출력(y) 쌍으로 된 데이터의 관계 학습



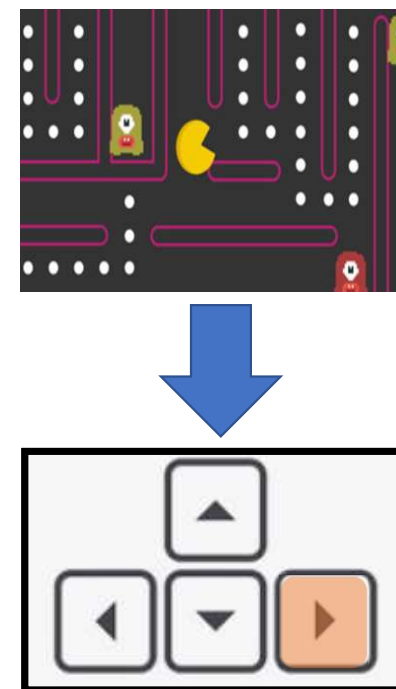
비지도 학습 (Unsupervised learning)

- 입력 (X) 데이터를 정해진 기준에 따라 관계 학습



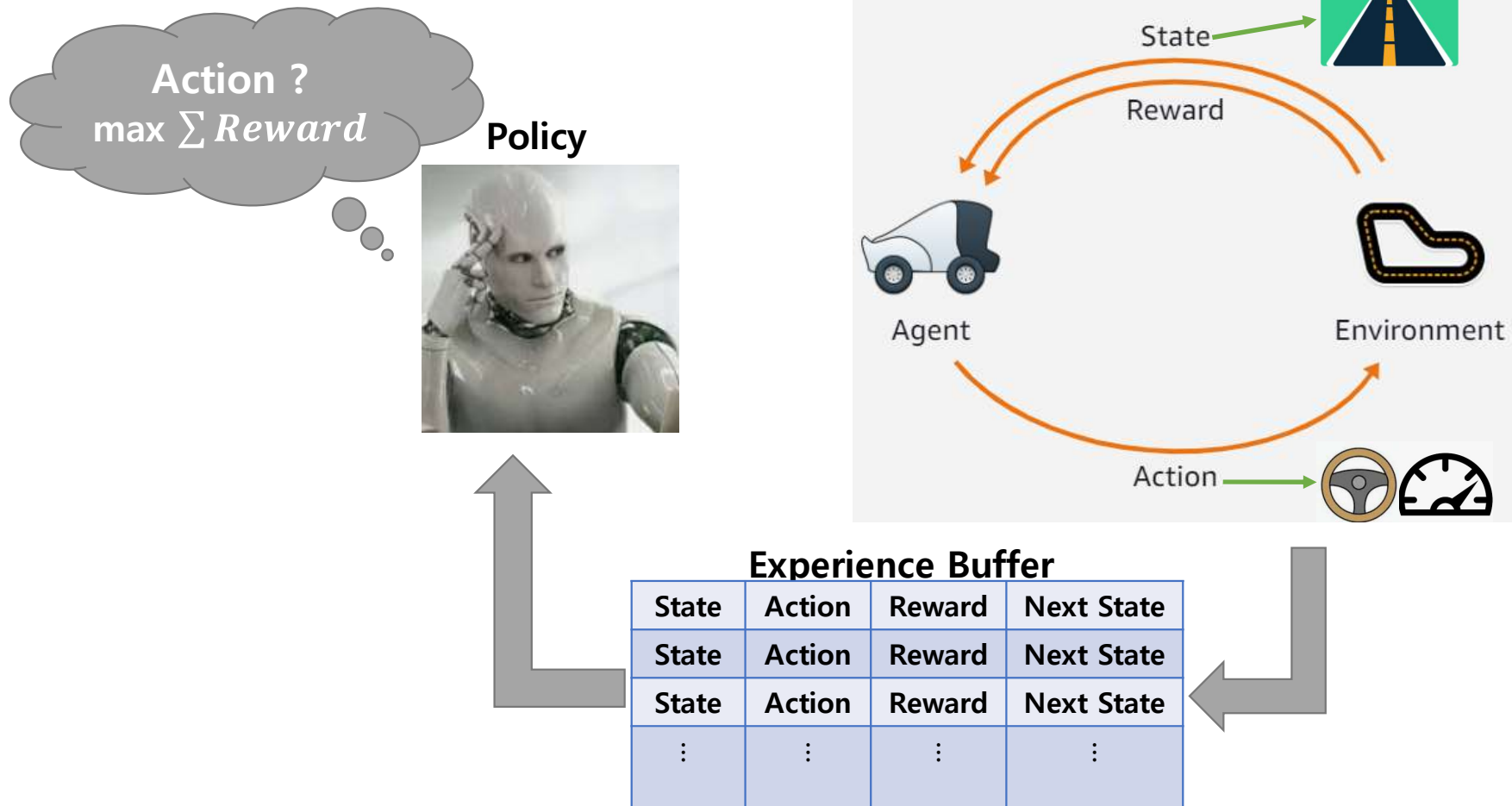
강화 학습 (Supervised learning)

- 보상의 합이 최대인 행동을 학습



2. Reinforcement Learning

2) 강화학습 프로세스



2. Reinforcement Learning

3) Discount Factor (할인계수)

- 미래에 받은 보상의 가치를 과거의 시점으로 환산한다고 하자. 이때, 이전 step 에 할인되는 비율을 Discount Factor 라 한다. (일반적으로 gamma, γ 로 표현하며 범위는 0과 1 사이)
- 현재 시점 t 에서 Action 또는 State 의 가치는 아래와 같이 계산된다.
(일반적으로 '할인된 미래 보상의 합' 또는 Return이라고 부른다.)

$$r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots = \sum_{i=t}^T \gamma^{i-t} r_i$$

(γ 는 할인계수)
(r_k 는 시점 k에서 받는 보상)

- 강화학습을 구체적으로 정의하면, 어떤 Action 또는 State 이 주어졌을 때 '할인된 미래 보상의 합' 이 어느 정도인지 측정하는 것을 학습하는 과정이다.
- 문제 상황에 따라 적절한 Discount Factor를 설정하는 일은 중요하다.
- Discount Factor 는 일반적으로 0.95 ~ 0.99 사이로 설정하며 연속적인 환경에서는 0.999 와 같이 1에 가깝게 설정하기도 한다. (why?)

2. Reinforcement Learning

4) Reward 희소성 문제

- 보상이 간헐적이면 강화학습 모델이 행동의 가치를 잘 학습하지 못할 수 있다. (why ?)
이것을 '보상의 희소성 문제' 라 한다.
- 이러한 보상의 희소성 문제를 해결하기 위한 방법 중 하나는
보상을 '단계화' 하여 나누어 주는 것이다.

관련 영상

(스키너 1) <https://www.youtube.com/watch?v=TtfQlkGwE2U>

(스키너 2) <https://www.youtube.com/watch?v=vGazyH6fQQ4>

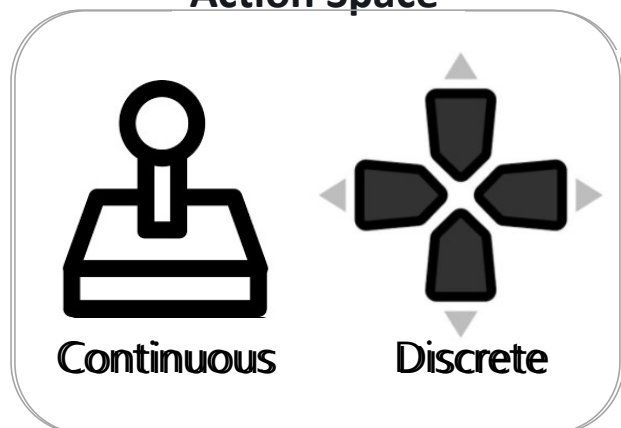
(강형욱) https://www.youtube.com/watch?v=wc_lhjhCfJo

2. Reinforcement Learning

5) DeepRacer 에서 제공하는 강화학습 알고리즘

	PPO	SAC
Action Space	Works in both discrete and continuous action spaces	Works in a continuous action space
Action Policy	On-policy	Off-policy
Exploration	Uses entropy regularization	Adds entropy to the maximization objective

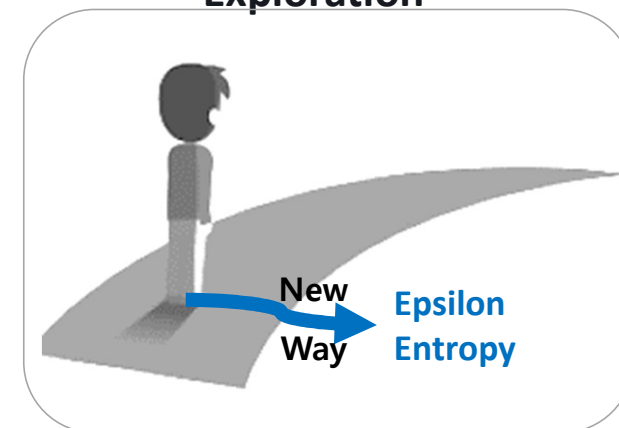
Action Space



Action Policy



Exploration



3. 온라인 환경에서의 DeepRacer 실습

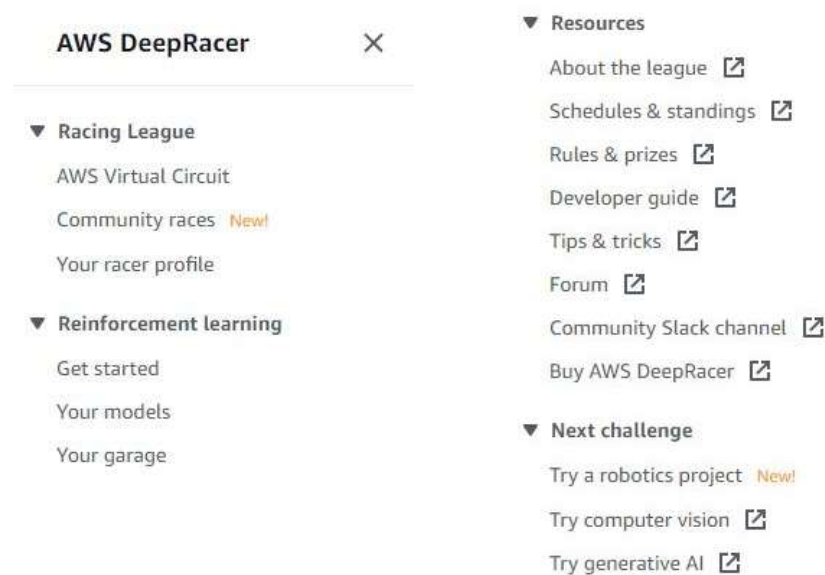
3. 온라인 환경에서의 DeepRacer 실습 – Console 접근

<https://console.aws.amazon.com/deepracer>

The screenshot displays the AWS DeepRacer console. The top navigation bar includes the AWS logo, a search bar, and links for PAI, the current region (Seoul), and support. A left-hand sidebar contains a menu with categories like 'Racing League', 'Reinforcement learning', 'Resources', and 'Next challenge'. The main content area features a large header with the text 'AWS DeepRacer' and 'The fastest way to get rolling with machine learning', accompanied by a 'Create model' button. Below this, there's a 'What is it' section with a video player showing a race car. To the right, a 'DeepRacer League' section provides an overview and links to the Virtual Circuit. A 'Pricing (US)' section is also visible at the bottom right. The footer contains copyright information and links to privacy, terms, and help.

3. 온라인 환경에서의 DeepRacer 실습 – Menu Guide

- AWS DeepRacer Console의 메뉴는 크게 Racing League, Reinforcement learning, Resources, Next challenge로 나눌 수 있습니다.



3. 온라인 환경에서의 DeepRacer 실습 - 모델 훈련

- 1. 메뉴에서 Reinforcement learning - Get started를 클릭합니다.

▼ Reinforcement learning

Get started

Your models

Your garage

3. 온라인 환경에서의 DeepRacer 실습 – 모델 훈련

- 2. Step 2: Create a model에서 Create model 버튼을 클릭합니다.

Step 2: Create a model

Simply follow the steps in the console to build, train and evaluate your model and enter the AWS DeepRacer League. With [AWS Free Tier](#), you will receive 10 free hours to train or evaluate models and 5GB of free storage during your first month. This is enough to train your first time-trial model, evaluate it, tune it, and then enter it into the AWS DeepRacer League. This offer is valid for 30 days after you have used the service for the first time.

Create model

3. 온라인 환경에서의 DeepRacer 실습 – 모델 훈련

- 3. Training details에서 Model name과 description을 작성합니다.

Training details

Model name

The model name must be unique and can have up to 64 characters. Valid characters are a-z, A-Z, 0-9, and - (hyphen). No spaces or underscores.

Training job description - optional

The model description can have up to 255 characters.

3. 온라인 환경에서의 DeepRacer 실습 - 모델 훈련


- 4. Environment simulation에서 훈련에 사용할 트랙을 선택합니다.

Environment simulation Info

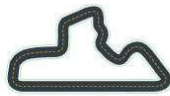
Simulated environment emulates a track to train your model.

Choose a track

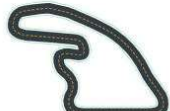
☐ **Baja Highway**
At 65.18m, The Baja Highway extends upon the shorter turnpike with pro level difficulty and even more opportunities for generating speed. A second massive straightaway mid course is similarly capped by a hairpin and increasing radius turn, along with an unforgiving technical section full of wrought with tight cutbacks and hairpins.




☐ **Baja Turnpike**
At 40.79m, the Baja Turnpike is a short angular track that mixes full throttle straightaways with hard decelerating turns. Signature features include a long high speed sprint over the finish line end capped by a harsh hairpin and tight 90 degree turn. While you won't find any sand traps on this course there are no shortage of places to get stuck.




☐ **Kuei Raceway**
Kuei Raceway is a fast short track (46.15m) with a friendly high speed arch and two straightaways interspersed between 2 hairpins, a chicane, and a technical cutback hairpin. It is named after 3rd place DeepRacer League finalists Kuei of NCTU CGI Taiwan.




☐ **Cosmic Circuit**
Cosmic Circuit expands the universe of its short loop cohort (62.24m). The pro difficulty course adds a deceptively technical section with multiple 90 degree turns separating meandering sprints void of a true straightaway.



☐ **Cosmic Loop**
Cosmic Loop is a short track (46.18m) that features a single sprint straightaway and a challenging technical section with multiple 90 degree turns and two hairpins. A successful run will require skillful command of time (steps) and (action) space.



☐ **Kuei Super Raceway**
Kuei Super Raceway adds pro difficulty to its short track counterpart by extending overall length in the form of a straightaway with a technical 4x double apex section. It is named after 3rd place DeepRacer League finalists Kuei of NCTU CGI Taiwan.





3. 온라인 환경에서의 DeepRacer 실습 – 모델 훈련


- 5. Race type에서 훈련할 race type을 선택합니다.
 - Race type은 Time trial, Object avoidance, Head-to-head racing 중 하나를 선택할 수 있습니다.

Race type

Choose a race type

☒ **Time trial**
The agent races against the clock on a well-marked track without stationary obstacles or moving competitors.


☐ **Object avoidance**
The vehicle races on a two-lane track with a fixed number of stationary obstacles placed along the track.


☐ **Head-to-head racing**
The vehicle races against other moving vehicles on a two-lane track.


3. 온라인 환경에서의 DeepRacer 실습 – 모델 훈련

- 6. Training algorithm and hyperparameters에서 훈련 알고리즘과 hyperparameter를 설정합니다.

Training algorithm and hyperparameters [Info](#)

☒ PPO
A state-of-the-art policy gradient algorithm which uses two neural networks during training – a policy network and a value network.

☐ SAC
Not limiting itself to seeking only the maximum of lifetime rewards, this algorithm embraces exploration, incentivizing entropy in its pursuit of optimal policy.

▼ Hyperparameters

Gradient descent batch size

- ☐ 32
☒ 64
☐ 128
☐ 256
☐ 512

Number of epochs

10

Integer between 3 and 10.

Learning rate

0.0003

Real number between 0.00000001 (1e-8) and 0.001 (1e-3).

Entropy

0.01

Real number between 0 and 1.

Discount factor

0.999

Real number between 0 and 1.

Loss type

- ☐ Mean squared error
☒ Huber

Number of experience episodes between each policy-updating iteration

20

Integer between 5 and 100.

3. 온라인 환경에서의 DeepRacer 실습 – 모델 훈련

- 6. Define continuous action space 에서 행동 공간에 대해 정의합니다.

- **PPO** 알고리즘의 경우 Continuous action space 와 Discrete action space 중 하나를 선택할 수 있습니다.

Define continuous action space [Info](#)

In a continuous action space setting, the agent learns to pick the optimal speed and steering values from the min/max bounds you provide through training. Providing a range of values for the model to pick from seems to be the better option but the agent has to train longer to learn to choose the optimal actions.

Steering angle

The steering angle determines the range of steering angles in which the front wheels of your agent can turn.

Left steering angle range
 degrees
Values are between 0 and 30.

Right steering angle range
 degrees
Values are between -30 and 0.

Speed

The speed determines how fast your agent can drive.
Min/max speed defines the range of speeds available to the agent while training.

Minimum speed
 m/s
Values are between 0.1 and 4.

Maximum speed
 m/s
Values are between 0.1 and 4.

[Reset to default values](#)

Dynamic sector graph

The graph shows a semi-circle representing the steering angle range from -30 to 30 degrees. The radius represents the speed range from 0.1 to 4 m/s. A blue shaded sector indicates the current selected range. A car icon is at the center. A legend at the bottom says: Select and drag an arrow to change the steering angle and speed.

[Cancel](#) [Previous](#) [Next](#)

3. 온라인 환경에서의 DeepRacer 실습 – 모델 훈련


- 7. Choose vehicle shell and sensor configuration 에서 agent를 선택합니다.

Choose vehicle shell and sensor configuration


Vehicle shell with sensor configuration (2)

Q Search < 1 >

☒ kakaka
Sensor(s): Camera; Shell: DeepRacer



☐ The Original DeepRacer
Sensor(s): Camera; Shell: DeepRacer



Cancel Previous Next

3. 온라인 환경에서의 DeepRacer 실습 – 모델 훈련

- 8. Reward function에서 보상함수를 작성합니다.
- Reward function examples에서 보상함수 샘플을 선택할 수 있습니다.

Reward function [Info](#)

The reward function describes immediate feedback (as a score for reward or penalty) when the vehicle takes an action to move from a given position on the track to a new position. Its purpose is to encourage the vehicle to make moves along the track to reach its destination quickly. The model training process will attempt to find a policy which maximizes the average total reward the vehicle experiences.

Code editor

Reward function examples

Reset

Validate

```
1 def reward_function(params):
2     """
3     Example of rewarding the agent to follow center line
4     """
5
6     # Read input parameters
7     track_width = params['track_width']
8     distance_from_center = params['distance_from_center']
9
10    # Calculate 3 markers that are at varying distances away from the center line
11    marker_1 = 0.1 * track_width
12    marker_2 = 0.25 * track_width
13    marker_3 = 0.5 * track_width
14
15    # Give higher reward if the car is closer to center line and vice versa
16    if distance_from_center <= marker_1:
17        reward = 1.0
18    elif distance_from_center <= marker_2:
19        reward = 0.5
20    elif distance_from_center <= marker_3:
21        reward = 0.1
22    else:
23        reward = 1e-3 # likely crashed/ close to off track
24
25    return float(reward)
```

3. 온라인 환경에서의 DeepRacer 실습 – 모델 훈련

- 9. Stop conditions에서 학습의 Maximum time을 설정합니다.

Stop conditions [Info](#)

Set the conditions for your training job to stop. To avoid run-away jobs, you can limit the length of a job to within a maximum time period (**Maximum time**).

The training will stop when the specified criteria is met. When your model has stopped training, you will be able to clone your model to start training again using new parameters.

Maximum time

Maximum time must be between 5 and 1440 minutes.

3. 온라인 환경에서의 DeepRacer 실습 – 모델 훈련

- 10. Automatically submit to the DeepRacer League에서 리그 참여 여부를 선택합니다.

Automatically submit to the DeepRacer League

☒ Submit this model to the DeepRacer League automatically after training completion and get a chance to win prizes.

Select DeepRacer League leaderboard to submit model to

July Qualifier (Time trial) ▼

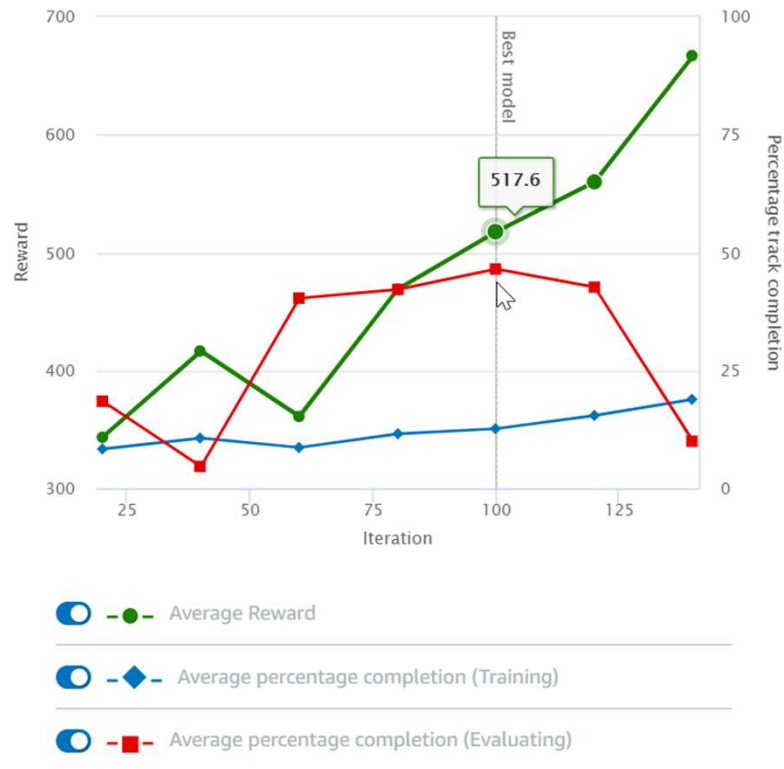
There are no fees or costs associated with submitting a model to the DeepRacer League.

- 리그에 참여하지 않아도 훈련을 진행할 수 있지만 참여하지 않는다면 모델 훈련 후 평가 시뮬레이션을 진행할 수 없습니다.

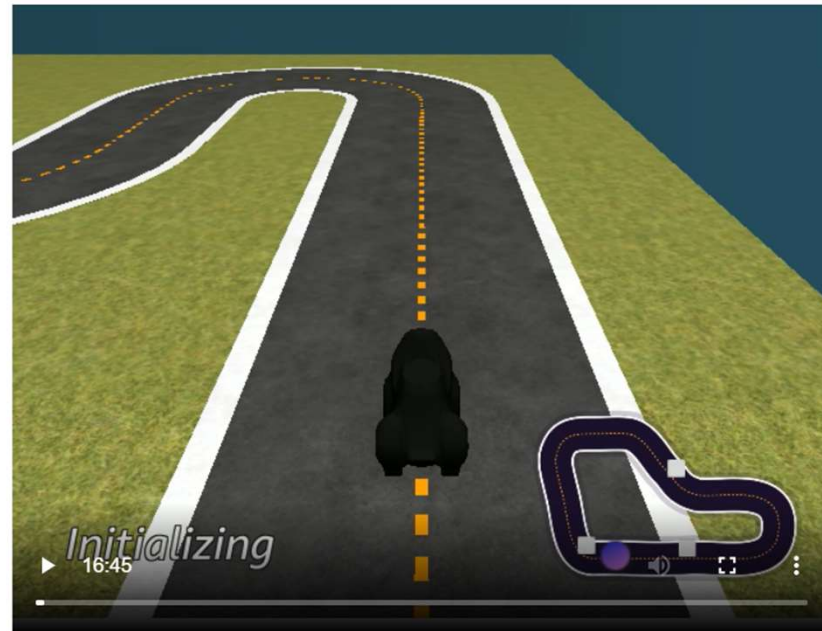
3. 온라인 환경에서의 DeepRacer 실습 – 모델 훈련

- 11. Reward graph and Simulation video stream에서 훈련 상황을 확인합니다.

Reward graph [Info](#)



Simulation video stream



3. 온라인 환경에서의 DeepRacer 실습 - 모델 평가 시뮬레이션

- 1. 메뉴에서 Reinforcement learning - Your models를 클릭합니다.

▼ Reinforcement learning

Get started

Your models

Your garage

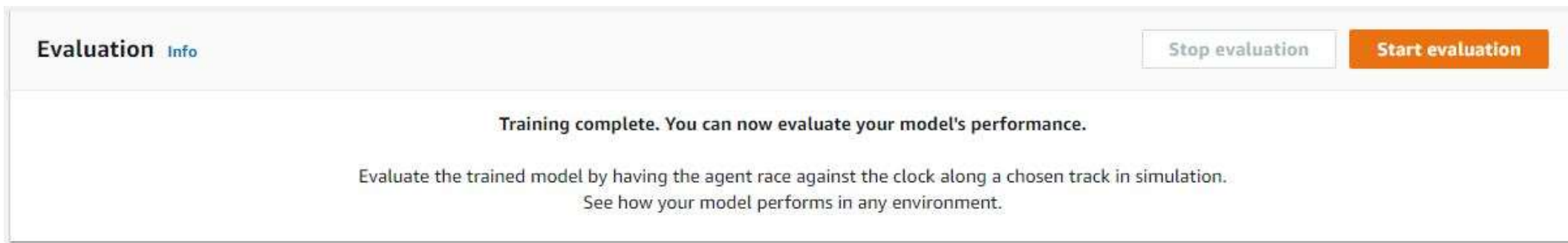
3. 온라인 환경에서의 DeepRacer 실습 – 모델 평가 시뮬레이션

- 2. 평가할 모델을 선택합니다.

	Name ▼	Description ▼	Status ▼	Algorithm ▼	Sensors ▼
<input type="radio"/>	pai-test2		Ready	PPO	Camera
<input type="radio"/>	pai-test		Ready	PPO	Camera
<input type="radio"/>	test		Ready	PPO	Camera

3. 온라인 환경에서의 DeepRacer 실습 – 모델 평가 시뮬레이션

- 3. Evaluation에서 Start evaluation 버튼을 클릭합니다.



3. 온라인 환경에서의 DeepRacer 실습 – 모델 평가 시뮬레이션

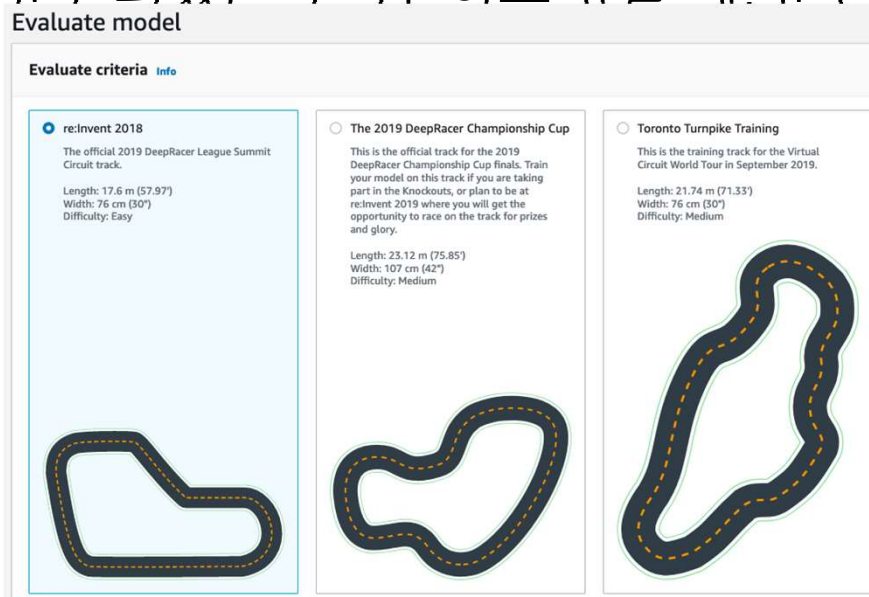
- 4. Evaluate model에서 평가에 사용될 트랙을 선택합니다.

- 훈련할 때 사용한 트랙과 동일하거나 비슷한 것을 선택해야 좋은 평가를 받을 수 있지만 아무 트랙이나 선택해도 평가가 가능합니다.

- 모델이 general하게 훈련되었는지 확인하고 싶을 때 훈련 시 사용한 트랙이 아닌 다른 트랙을

랙을

선택합니다.



3. 온라인 환경에서의 DeepRacer 실습 – 모델 평가 시뮬레이션

- 5. Race type에서 racing type을 선택합니다.

- 훈련에서 사용된 racing type과 동일한 것을 선택해야 좋은 평가를 받을 수 있지만
다른

racing type을 선택해도 평가가 가능합니다.

- 모델이 general하게 하려디어느기 하이찬고 시으 때 하려 시 사요하 racing type이
아닌 다른

Race type

Choose a racing type

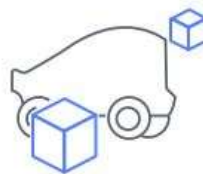
☒ Time trial

The agent races against the clock on a well-marked track without stationary obstacles or moving competitors.



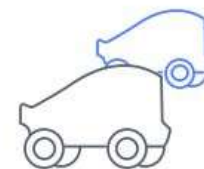
☐ Object avoidance

The vehicle races on a two-lane track with a fixed number of stationary obstacles placed along the track.



☐ Head-to-head racing

The vehicle races against other moving vehicles on a two-lane track.



3. 온라인 환경에서의 DeepRacer 실습 – 모델 평가 시뮬레이션

- 6. Virtual Race Submission에서 모델 제출 여부를 선택합니다.

Virtual race submission

Virtual races [Info](#)

Congratulations training your model, now see how your model stacks up. Submit your model to participate in the virtual race. Your model will be ranked based on the average time it takes to complete a lap on the race track. Your results will be displayed on the leaderboard. Win prizes, no fees or costs for entering the virtual league and unlimited race submissions.

☐ Submit model after evaluation
Win prizes, no fees or costs for submitting a model to the virtual league.

3. 온라인 환경에서의 DeepRacer 실습 – 모델 평가 시뮬레이션

- 7. Simulation에서 평가가 진행되는 것을 확인합니다.

- Simulation video stream에서 평가 시뮬레이션을 확인할 수 있습니다.
- Evaluation results에서 완주 소요 시간 및 완주율, 완주여부를 확인할 수 있습니다.
- 평가가 끝나기 전에 Stop evaluation 버튼을 통해 평가를 정지할 수 있습니다.

Evaluation [Info](#) [Download logs](#) [Stop evaluation](#) [Start new evaluation](#)

Simulation video stream


test
2/3
Race | 00:37.448
Reset | 0
00.82 m/s
Time trial evaluation

Evaluation results

Trial	Time (MM:SS.mmm)	Trial results (% track completed)	Status
1	00:33.656	100%	Lap complete

3. 온라인 환경에서의 DeepRacer 실습 – 모델 평가 시뮬레이션

- 8. 평가 결과를 확인합니다.


- Evaluation results에서 모든 주행 시도에 대한 결과를 확인할 수 있습니다.
- 평가가 완료된 후 평가 시뮬레이션은 더 이상 확인할 수 없습니다.
- 모델 평가 결과가 좋지 않다면 훈련 시간, 보상 함수, hyperparameter 등을 재설정하는

방법으로 모델을

Evaluation [info](#)

[Download logs](#) [Stop evaluation](#) [Start new evaluation](#)

Simulation video stream



Simulation video stream not available.
Video is only available during evaluation.

Evaluation results

Trial	Time (MM:SS.mmm)	Trial results (% track completed)	Status
1	00:33.656	100%	Lap complete
2	00:34.728	100%	Lap complete
3	00:08.795	26%	Off track

3. 온라인 환경에서의 DeepRacer 실습 – Racing League

- **Racing League**는 두 가지로 나눌 수 있습니다.
 - AWS에서 후원하는 이벤트를 AWS DeepRacer League라고 하며 매월 개최됩니다.
처음에는 Open Division에 참여하게 되고 상위 10% 레이서는 Pro Division에 진출합니다.
 - 일반적인 사용자가 생성한 이벤트를 Community Race라고 합니다.

AWS DeepRacer League

Open Division

Pro Division

Community Race

Classic race

LIVE race

과제

과제!!

1. DeepRacer 콘솔 메뉴창에서 각 목록들을 클릭해서 들어가보고 어떤 내용들이 있는지 각각 3줄 이내로 요약
2. DeepRacer 온라인 환경에서 (기본 값으로) 모델을 훈련하기