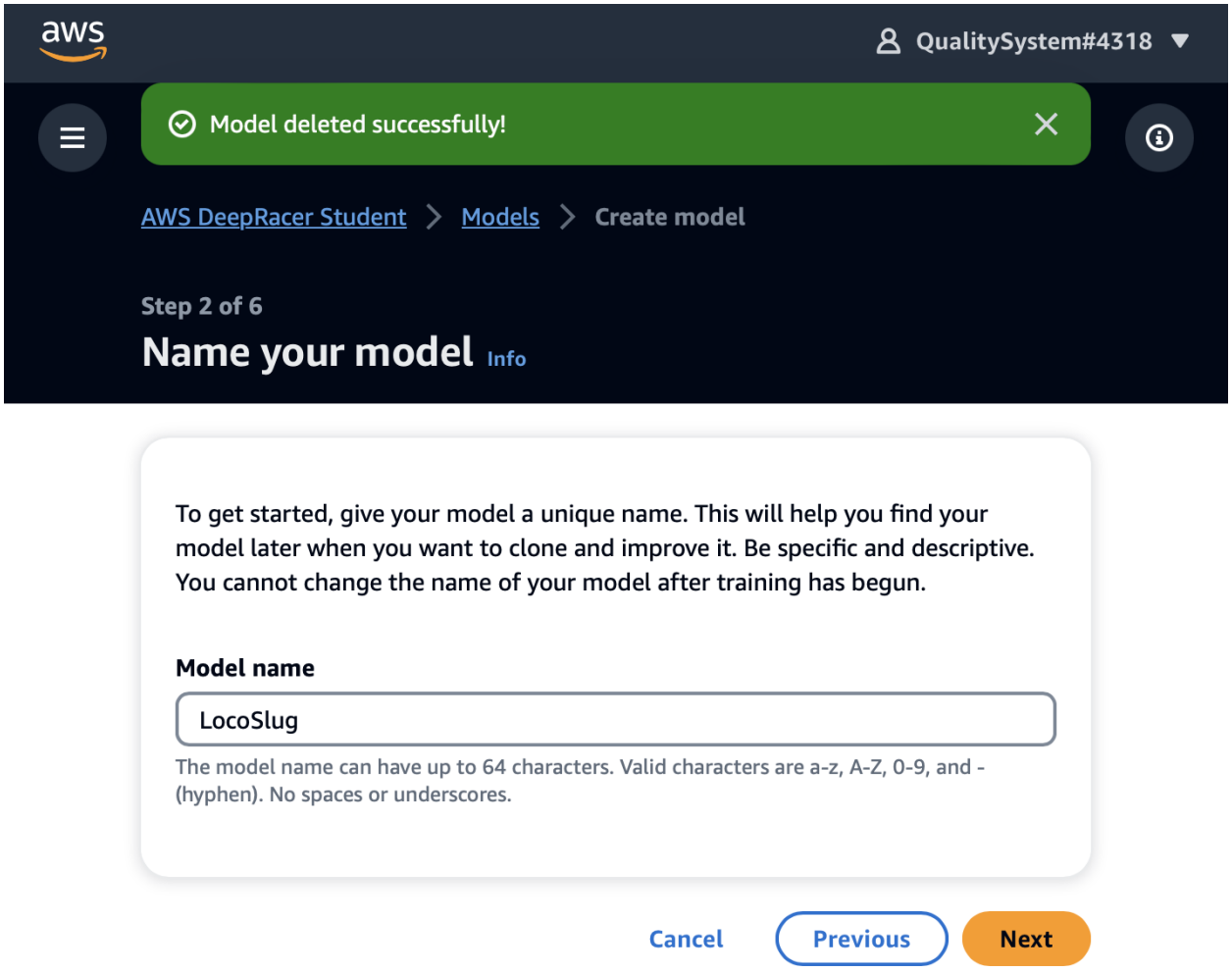
Lab 12

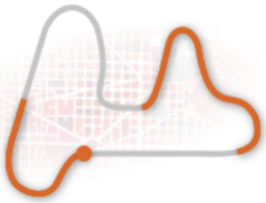
By **Vannarith Om**

Setting up the name of the first model.



Selecting re:Invent 2018 track the track to train the model

Po-Chun Speedway
Direction: Clockwise, Counterclockwise



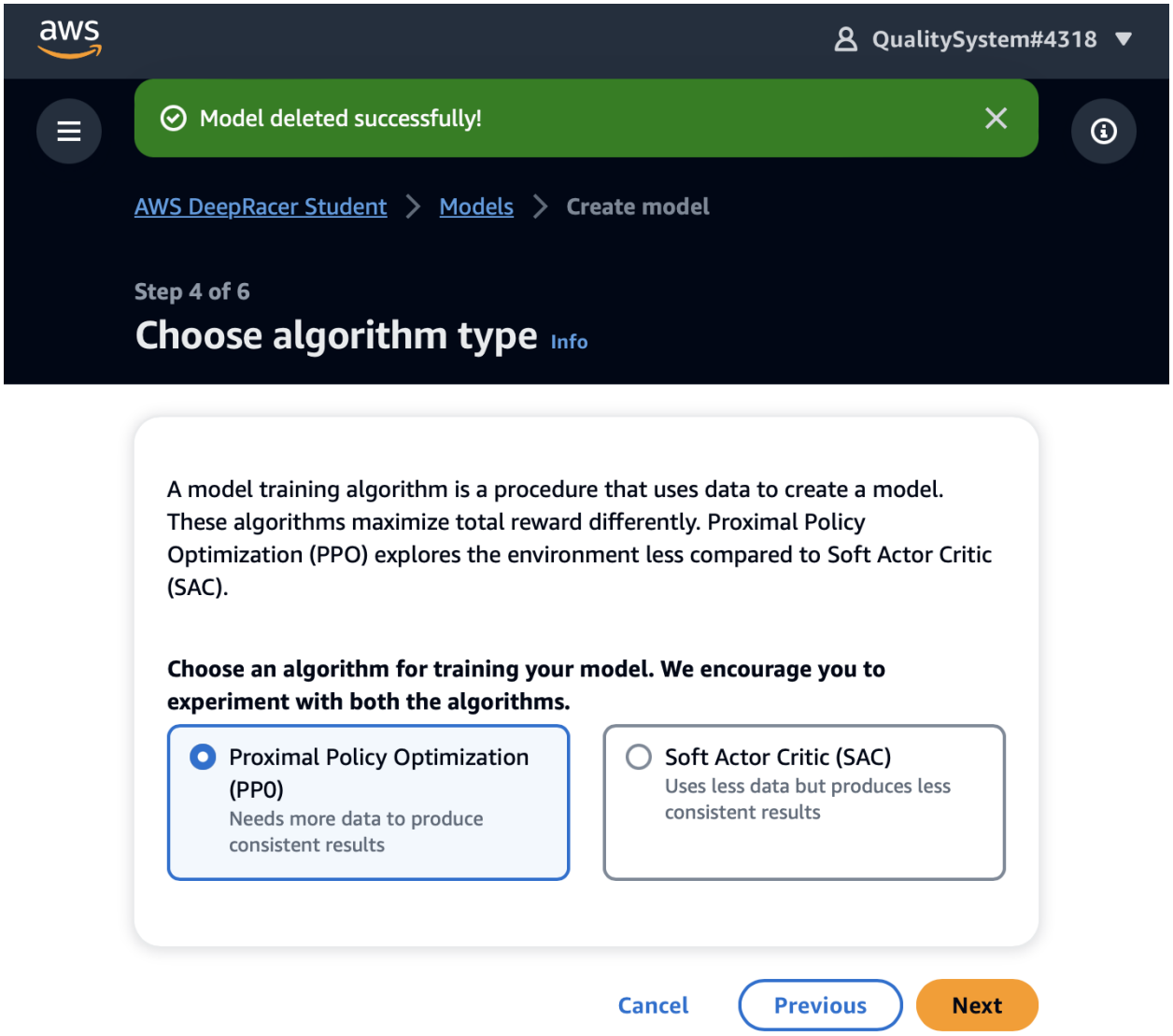
Smile Speedway
Direction: Clockwise, Counterclockwise

re:Invent 2018
Direction: Counterclockwise

A to Z Speedway
Direction: Clockwise, Counterclockwise

Cancel Previous Next

Selecting the the Proximal Policy Optimization (PPO) as the algorithm type.



Initial Reward system that balances following the middle line and maintaining the car inside the border lines. Confirming that the reward function can be validated.

```python
def reward_function(params):
    # Constants
    MAX_SPEED = 5.0
    MIN_SPEED = 2.0
    MAX_STEERING_ANGLE = 30.0

    # Read input parameters
    distance_from_center =
        params['distance_from_center']
    all_wheels_on_track =
        params['all_wheels_on_track']
    speed = params['speed']
    steering_angle = abs(params['steering_angle'])

    # Calculate reward
    reward = 0.0

    # Stay close to the centerline
    if distance_from_center < 0.1:
        reward += 1.0
    else:
        reward += 0.5

    # Stay within the track borders
    if all_wheels_on_track:
        reward += 1.0
    else:
        reward -= 1.0

    # Speed reward
    reward += min(speed/MAX_SPEED, 1.0)

    # Penalize steering too much
    if steering_angle > MAX_STEERING_ANGLE:
        reward -= 0.5

    return float(reward)
```

Python Ln 1, Col 1 Errors: 0 Warnings: 0

```python
    if distance_from_center < 0.1:
        reward += 1.0
    else:
        reward += 0.5

    # Stay within the track borders
    if all_wheels_on_track:
        reward += 1.0
    else:
        reward -= 1.0

    # Speed reward
    reward += min(speed/MAX_SPEED, 1.0)

    # Penalize steering too much
    if steering_angle > MAX_STEERING_ANGLE:
        reward -= 0.5

    return float(reward)
```

Python Ln 1, Col 1 Errors: 0 Warnings: 0

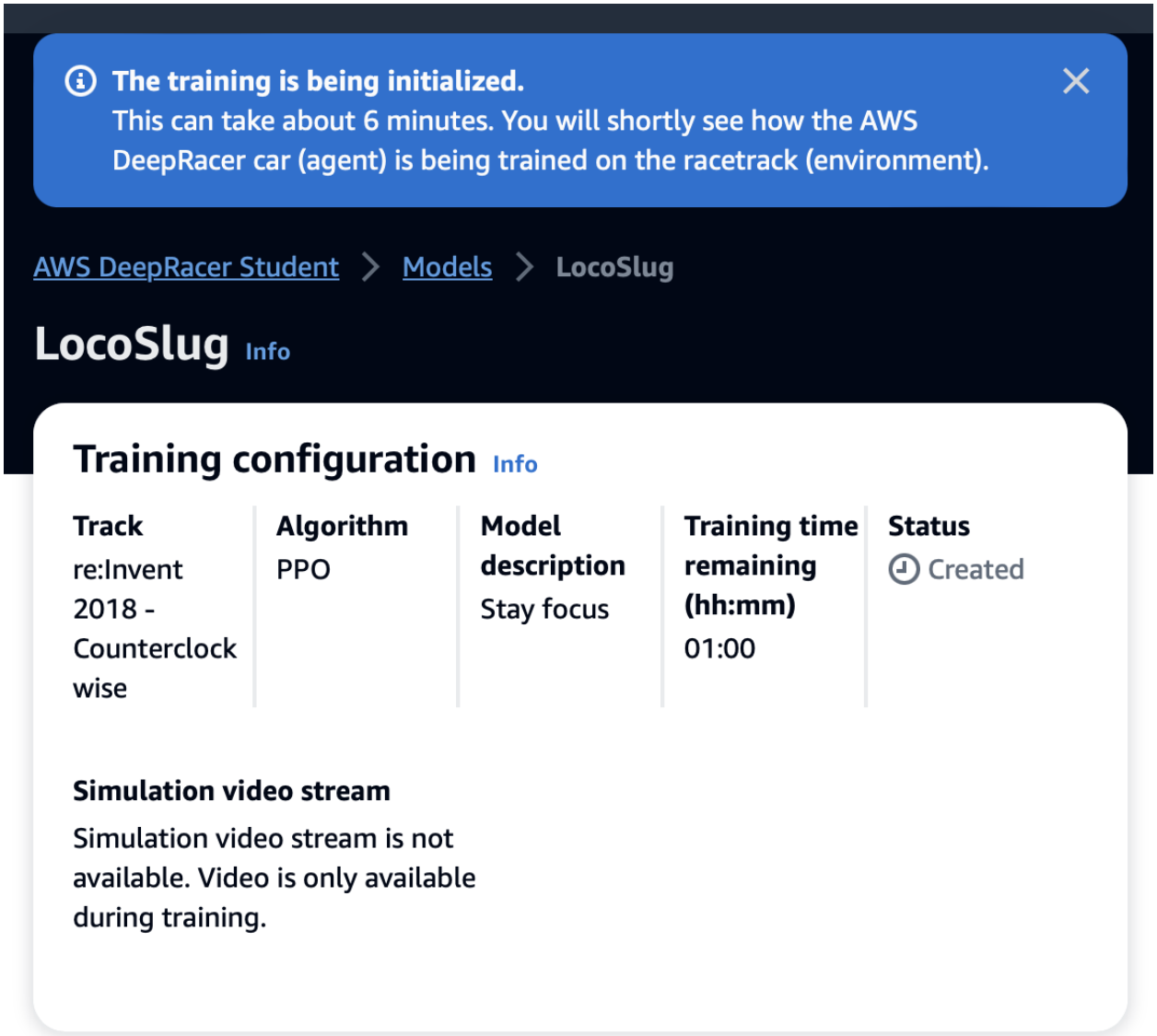
Reset Validate

Once you are done with changes, click Next.

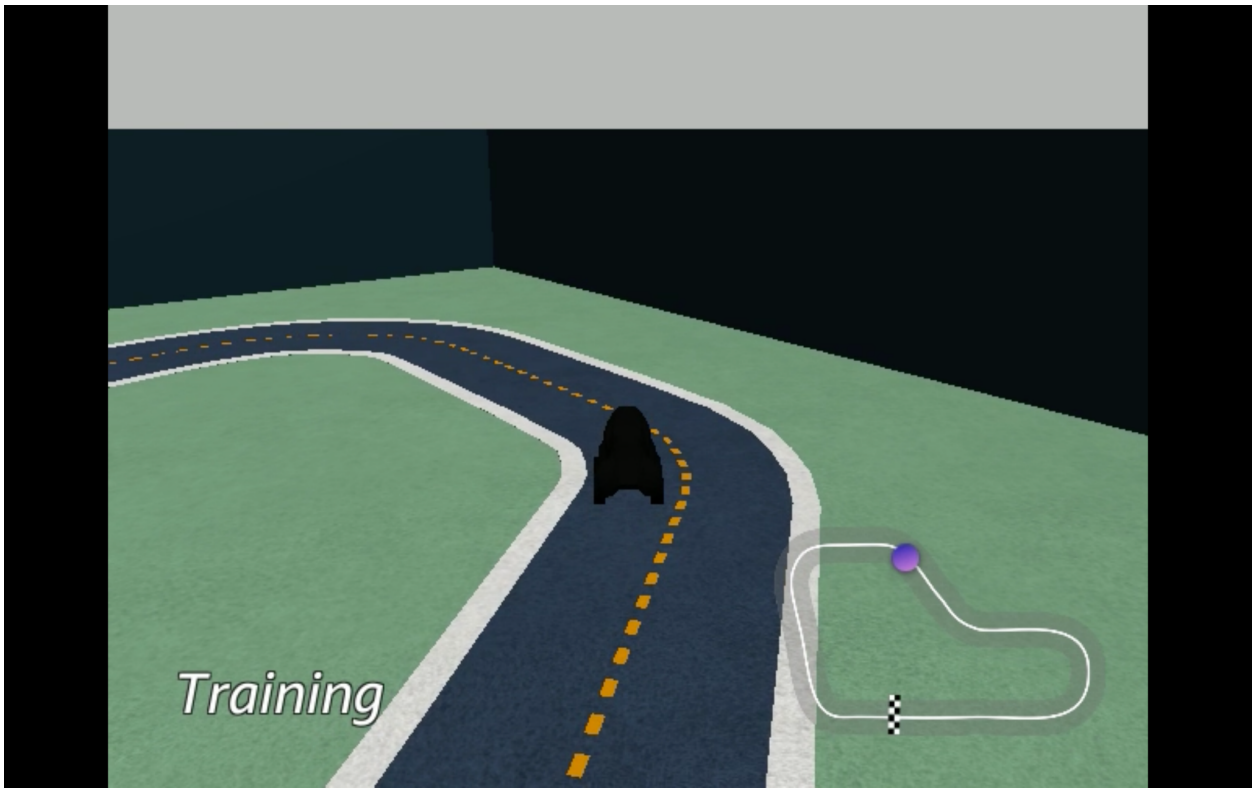
Reward function successfully validated!

Cancel Previous Next

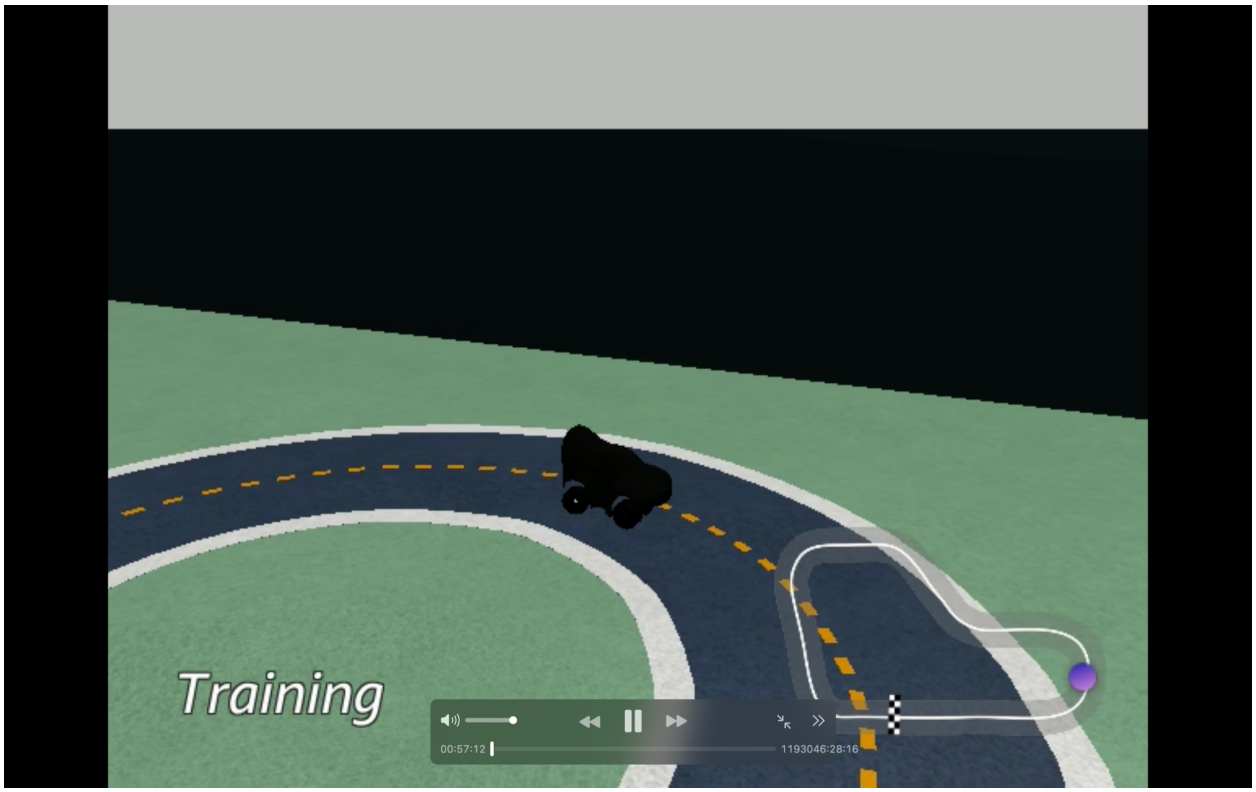
Setting the initial training to 60 minutes (one hour)



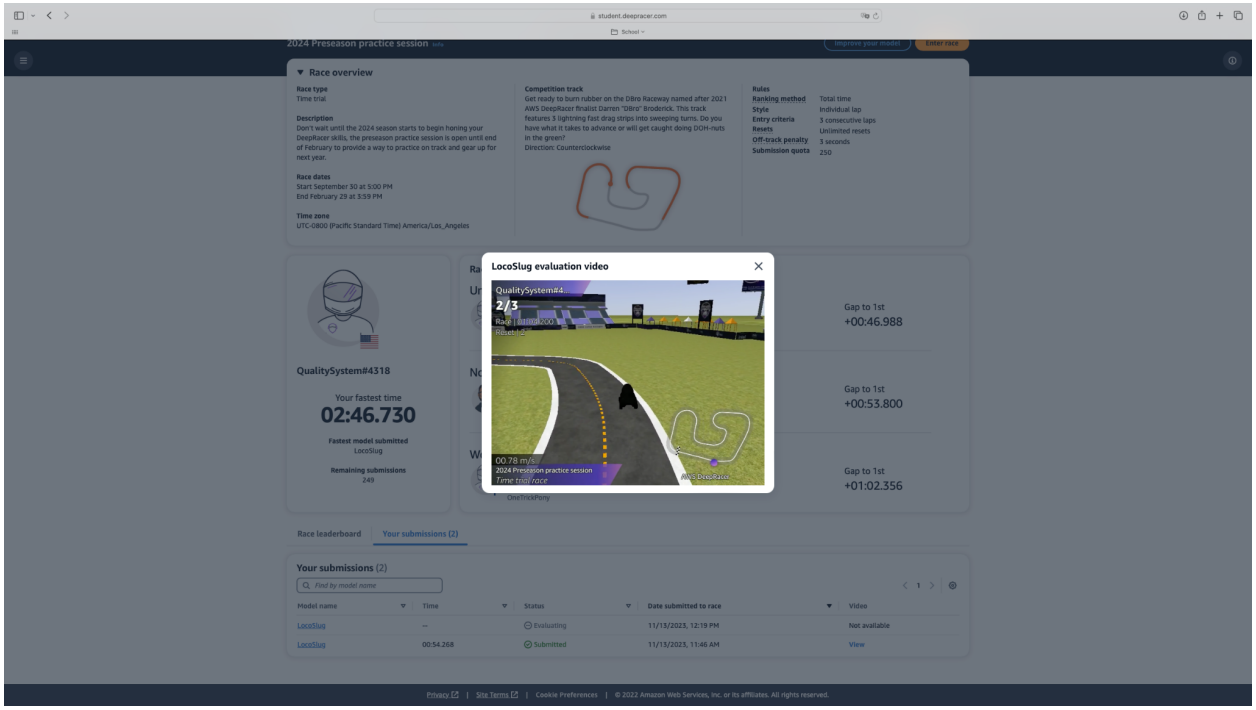
Car during first minutes of training process.



Car doing getting better at curves as time of training advances.



LocoSlug being evaluated for the preseason practice session



LocoSlug evaluation video

Copy of the original model with implementation of the Markov Decision Process in the rewarding funcion.