9회의 오픈 AI 세미나

1회 - python 설치 및 기본 개념 1

python을 설치하고, 코딩을 위한 환경을 설정합니다.

파이썬의 변수, 리스트, 조건문(if), 반복문(for, while)에 대한 내용을 배우고, 간단한 실습을 통해 내용을 익힙니다.

2회 - python 기본 개념 2

파이썬의 class를 이용해 직접 클래스를 만들고, 상속에 대한 개념을 배웁니다.

파이썬의 모듈을 직접 만들어보고, 자신의 모듈을 사용해봅니다.

외부 모듈을 직접 설치해보고 사용해봅니다.

3회 – 강화학습의 기본 개념

강화학습에서 state, action이 무엇인지 배우고, reward는 어떻게 계산하는 지를 배웁니다.

state value와 policy에 대한 개념을 익히고 직접 계산해봅니다.

DP를 통해 state value와 policy를 향상시키는 법을 배우고, 이를 이용하여 미로 예제를 직접 풀어봅니다.

4회 – Model free control

DP를 사용하면 오목과 같이 state가 많은 환경에서는 사용하는 것이 불가능합니다. 따라서 model의 형태에 상관없이 model을 학습하는 방법인 MC와 SARSA, Q-learning을 익히고 이를 활용하여 cliff walking 문제를 풀어봅니다.

5회 – neural network

앞으로의 모델은 nn을 활용해서 설계하게 됩니다. 따라서 neural network의 기본 원리인 gradient descent를 직접 구현해봅니다. 그 후, python의 모듈인 pytorch를 이용하여 gradient descent를 다 시 구현해봅니다.

그 후 pytorch로 neural network를 설계해보고, 간단한 문제를 해결해봅니다.

6회 – gym 튜토리얼

python의 gym은 강화학습을 위한 환경을 제공합니다. gym을 활용하는 법을 배우고, gym의 frozen lake 예제를 직접 해결해봅니다.

7회 – deep Q learning

Q-learning을 2개의 network를 사용하여 문제를 해결하는 기법을 DQN(deep Q-learning)이라 합니다. DQN을 이용하여 gym의 cartpole 문제를 풀어봅니다.

8회 – Policy gradient

지금까지는 state value를 계산하는 방식으로 문제를 해결했습니다. 이번에는 policy 자체를 최적화하는 알고리즘인 Actor Critic과 Q-value 대신 Advantage를 적용한 A2C(Advantage Actor Critic)에 대해 배우고, 이를 이용하여 Acrobot 문제를 풀어봅니다.

9회 – A2C on continuous space

할 수 있는 action의 수가 유한한 경우, softmax function을 통해 각 action을 취할 확률을 계산하였습니다. 그러나 action이 continuous space인 경우에는 gaussian distribution(normal distribution)을 사용하여 action을 결정해야 합니다.

그 외에도 학습을 효율적으로 진행시키기 위해 가중치 초기화 기법(xavier), state normalization을 적용해봅니다.