

제목 : Playing Atari with Deep Reinforcement Learning

1. 내용

딥러닝 모델이 고차원의 시각적 입력에서 직접적으로 제어 정책을 학습하는데 성공한 최초의 모델

Atari 2600 게임 환경에서 비디오 데이터를 사용하여 정책을 학습하는 방법을 제시

기존의 Q-Learning을 뛰어넘는 알고리즘을 소개했습니다. 이 알고리즘은 다음과 같은 이점이 있음

1. 데이터 효율성:

각 스텝에서의 경험은 많은 가중치 업데이트에 사용되며, 이는 뛰어난 데이터 효율성을 가능하게 함.

2. 상관 관계 해결:

연속적인 샘플로부터 직접적으로 학습하는 것은 상관 관계를 무너뜨리고 업데이트를 안정적으로 도와줌.

이 논문에서 사용된 Q 네트워크의 구조는 다음과 같음

1. 입력으로 84x84 크기의 이미지가 들어옴
2. 8x8 크기의 16개 커널이 합성곱을 진행 (stride=4).
3. 4x4 크기의 32개 커널이 합성곱을 진행 (stride=2).
4. 256개의 유닛을 가진 완전 연결층을 통과