|  |  |
| --- | --- |
| 과목명 | 강화학습 (10-3) |
| 이름 | 이덕수 |
| 학번 | 20161600 |

이름과 학번을 적고, 답변을 적어주세요.

실습: 10주차 수업 내용을 자유롭게 요약하세요(A4 1 page 분량, 넘어도 상관없음)

- GPT-1 이해하기

자연어 처리에는 학습 데이터가 필요 이때 학습 데이터는 2가지 학습 데이터가 필요함

1. 정답이 정해진 문제 (지도 학습에 사용)
2. 정답이 정해지지 않은 문제(비지도 학습에 사용)

**가령 감정 분석을 예를 들면**

감정분석을 하기 위해선 단어 사전에 기반한 긍정어 부정어 데이터 필요(지도학습)

데이터가 없다면 학습을 위해 데이터를 수집하고 라벨링 해야함

**가령 워드투 벡터를 예를 들면**

빈칸 앞뒤에 함께 등장하는 단어의 빈도를 알아볼 때 사람이 정답 데이터를 반드시 입력할 필요 없음(비지도 학습)

GPT-1은 여건이 된다면 사람이 라벨링 하는 것이 베스트 하지만 일일히 하기엔 너무 많은 작업량 = 단기간에 많은 데이터 생성 어려움

만약 라벨링을 하지 않으면 많은 데이터를 활용할수 있으나 정교함이 떨어짐

그러나 위의 두가지 과정을 함께 한다면 ? == 이것이 GPT-1

1. 매우 많은 데이터를 비지도 학습으로 학습 시켜 언어 모델 생성
2. 수행하는 작업에 맞게 라벨링을 한 데이터를 활용하여 지도학습 으로 추가 학습을 시킴(파인 튜닝)

GPT-1 > 파인튜닝 > 기존 모델에 다른 작업에 대한 추가 데이터를 활용 추가 훈련 으로 모델을 학습, 입력값 으로 제공하는 수준을 넘어 모델 가중치 자체를 매번 업데이트 해야함

GPT-2 > 제로샷 러닝 > 입력값으로 주어진 예시가 없고 해야하는 작업의 명세 만으로 다른 작업에 기존 모델을 재활용 하는 방법

GPT-3 > 퓨삿 러닝 > 기존 모델에 입력값 몇 개를 예시로 해서 작업을 수행하도록 학습

> 원샷 러닝 > 기존 모델에 하나의 입력값을 예시로 해서 다른 작업을 수행하도록 시키는 방법

GPT-2 는 세상에 공개 되지 못함 ??? 너무 무서운 기술이라 악용될 위험이 있어서 !!

또한 모델의 평가물을 사람이 직접 피드백 하여 성능을 향상시킴

이때 강화학습 이기에 보상이나 패널티를 정해야 하나 GPT-2 에서는 보상 및 패널티를 정하기 매우 어려움 그렇기에 = 사람이 직접 피드백 하여 평가 = 결과적 으로 더 우수한 언어 생성 모델 학습 가능!