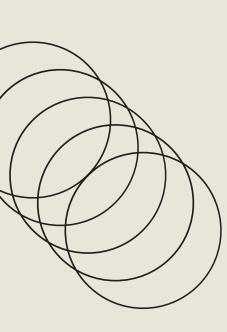


강한학습

712 PROJECT AISLA

20195237 장윤성 20215124 김수현



CONTENTS

01	02	03 —	04
프로젝트 목표	팀원 소개	서론	관련 연구 및 기술 동향
		문제설명, 딥러닝, 강화학습	딥러닝, 강화학습

 05
 06
 07

 방법
 실험
 기대효과

 강화학습시스템 정의
 실험환경, 평가방법
 프로젝트 결과물의 활용 예

01. 프로젝트 목표

각 사용자에게 맞는 개인화된 음식 추천을 위한 강화학습 알고리즘 개발

02. 팀원 소개

팀장: 빅데이터학과 20195237 장윤성

팀원: 빅데이터학과 20215124 김수현

03. 서론

문제 설명 :

메뉴 선택을 쉽게 하지 못하는 사람들을 위해 생각해낸 프로젝트입니다.

딥러닝으로 해결이 어려운 이유:

개개인의 음식 취향이 다르고 당장의 상황에 따라 데이터가 동적으로 달라지기 때문에 딥러닝으로 학습하기 어렵습니다.

강화학습으로 도전하는 이유:

각 사용자의 기분과 선호도가 달라 동적으로 변하기 때문에 레이블이 정해진 딥러닝으로 학습이 어렵고, 에이전트가 환경과 상호작용하며 연속적으로 학습할 수 있습니다.

또한, 사용자로부터 얻은 피드백을 보상으로 설정하여 에이전트가 사용자의 선호를 더 잘 이해하고 예측할 수 있게 됩니다.

03. 서론 - 학습 과정



초기학습:

보편적으로 인정되는 메뉴로 REINFORCE 알고리즘을 사용하여 강화학습 모델을 초기 학습합니다.

배포:

초기 학습된 모델을 사용자에게 배포하여 사용자의 피드백을 수잡합니다.

개인화된 학습:

사용자 피드백을 수집하여 각 사용자에게 맞는 음식 추천을 위해 모델 파라미터를 업데이트합니다.

세트별 업데이트:

아침, 점심, 저녁, 야식으로 구분된 4묶음 세트를 사용하여 모델을 업데이트합니다.

평가:

사용자 피드백에 따라 모델의 성능을 평가합니다.

04. 관련 연구 및 기술 동향

딥러닝으로 시도한 사례들:

[음식 추천 시스템]

많은 음식 추천 시스템들이 사용자의 클릭이나 구매 기록을 기반으로 CNN, RNN을 사용해 추천을 해왔습니다. 하지만 이 방식은 사용자의 변화하는 기분이나 감정을 반영하기 어렵습니다.

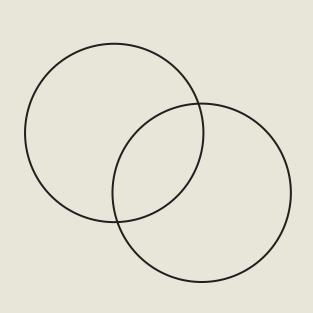
강화학습으로 시도한 사례들:

[개인화 추천 시스템]

Netflix, YouTube 등이 사용하는 강화학습 기반 추천 시스템은 사용자의 상호작용을 통해 지속적으로 학습하고, 사용자 경험을 개선합니다.

05. 방법

강화학습 시스템 정의

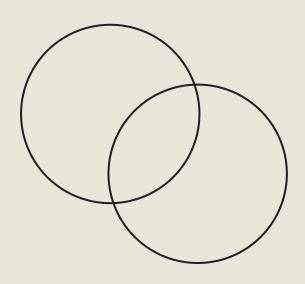


Agent	메뉴를 추천하는 주체로, 사용자의 기분, 이전에 선택한 메뉴, 현재 상태 등을 고려하여 다음 메뉴를 추천합니다.	
Environment	사용자의 기분, 선택 가능한 메뉴 리스트, 시간 등이 포함된 환경입니다. 사용자의 기분에 영향을 미치는 요인들 또한 환 경이 됩니다.	
State	Agent가 결정을 내리는 데 사용하는 정보로, 사용자의 현 재 기분, 이전 메뉴 선택, 시간 등으로 구성됩니다.	
Action	Agent가 할 수 있는 행동으로, 특정 메뉴를 추천하는 것입니다.	
	사용자의 선택으로 계산되며, Agent의 추천이 얼마나 좋은 지 평가하는 지표입니다. 만족독 높을수록 더 큰 보상을 받 습니다.	
강화학습 시스템 정의	Policy Network	
정책기반 강화학습 알고리즘 (pseudo code)	뒷 페이지에 짧은 의사코드 있습니다.	

05. 방법

정책기반 강화학습 알고리즘 (pseudo code)

의사코드



```
def reinforce(policy_network, optimizer, episodes):
  for episode in range(episodes):
    state = env.reset()
    done = False
    while not done:
      action_probs = policy_network(state)
      action = select_action(action_probs)
      next_state, reward, done = env.step(action)
     optimizer.zero_grad()
      loss = -torch.log(action_probs[action]) * reward
      loss.backward()
      optimizer.step()
     state = next_state
```



06. 실험



2024.05.21. 오후 03:35 최종 수정 강화학습을 위한 개인별 선호 음식 데이터 수집 설문기간 2024.05.21. 오후 03:13 ~ 제한 없음 📆 + 질문 추가 🛅 페이지 추가 *1. 오늘 기분을 골라주세요. 1점 최악, 10점 최고 1 2 3 4 5 6 7 8 별로다 이 이 이 이 이 종다 *2. 현재의 날씨를 골라주세요. 현재 기입된 날씨가 어느정도 겹치는 경우 비슷한 날씨끼리 복수선택이 가능합니다. 마음 실험 환경

Intel Core i5-8250U CPU google colab T4 GPU

데이터 수집:

네이버 폼을 활용하여 사용자의 기분, 날씨, 시간, 배고픔 정도, 인원 수, 성별, 나이대 등의 정보 수집

학습 환경:

강화 학습 모델에 입력으로 사용될 상태 정보 및 보상 정보 정의

알고리즘 선택:

REINFORCE 알고리즘 선택

보상함수 설계:

사용자 피드백을 반영하여 보상을 조정.

사용자가 좋아하는 음식에 대한 평가는 높은 보상을 부여하고, 만족하지 못한 음식에 대한 평가는 낮은 보상을 부여 (별점 1~5 까지 설정 예정)



06. 실험





1. 사용자 만족도 조사

직접 피드백: 사용자로부터 직접 받은 피드백 점수를 분석하여 평가합니다.

설문 조사 : 사용자 경험에 대한 설문 조사를 통해 모델에 대한 만족도를 추가로 평가합니다.

2. 반복적 학습 및 평가

학습 곡선 : 각 Epoch마다 모델의 성능 변화를 추적합니다.

피드백 루프: 사용자 피드백을 모델에 반영하고, 그 결과를 다시 평가합니다.

3. 평가 지표 - 정확성

평균 보상: 사용자가 제공한 피드백 점수의 평균을 측정합니다. 이는 모델이 사용자의 기대에 얼마나 부합하는지를 간단하게 보여줍니다.

정확률 (Precision): 추천 시스템이 '좋음'으로 예측한 항목 중 실제로 사용자가 좋아하는 항목의 비율입니다.

07. 기대효과

프로젝트 결과물의 활용 예

- 1. 음식점 및 카페 추천
- 2. 미디어 콘텐츠 추천 시스템
- 3. 사용자의 니즈를 파악한 개인별 마케팅 자료 제공
- 4. 건강 관리 식단 어플

THANK YOU