



Reinforcement learning Design Study

🕒 Created	@Mar 24, 2020 1:15 PM
🏷️ Tags	

Week8 Reinforcement Learning assignment 2

[투빅스 13기 @JM Heo]

[과제 2]

강화학습 설계해보기!

실생활에서 강화학습으로 해결 할만한 주제를 생각해보고, environment, state, reward, action 을 정의해주세요.


실제로 강화학습이 어떻게 적용될 수 있는지 한번쯤은 고민해주시길 바라는 마음으로 내는 과제이니 가볍게 해주세요.

물론 구체적으로, 열심히 해주시면 더 좋습니다. 제가 생각해보지 못했던 신선하고 기발한 주제면 더 좋아요!

저도 매우 부족하지만 열심히 피드백 해드리겠습니다.

예)

주제 : 강화학습으로 교차로 신호등 제어하기

Environment : 교차로 

State : 각 차선마다 대기중인 차의 개수, 가장 오래 기다린 차의 대기시간, 대기중인 차들의 평균 대기 시간

Action : 신호등 신호

Reward : (이번 step에 내보낸 차의 대기 시간) - (해당 교차로에서 가장 오래 기다린 차의 대기 시간) * 0.5



1. 주제

주제 도출과정

나의 관심 분야

- Domain
 - E-commerce
 - Digital Marketing
 - Omni Channel 분석
 - Social Network Services
-
- 인터넷을 통한 인구 활동 증가
 - 온라인 쇼핑 증가로 인한 오프라인 매장 수요 감소
 - 코로나로 인한 쿠팡 등 온라인 쇼핑몰의 수요 급증
-
- 온라인 쇼핑몰에서 사람들은 어떠한 행동패턴을 보일까?
 - 어떠한 행동패턴을 조사해야할까?
 - 행동패턴을 조사하는 방법은?
-

2. 강화학습 적용

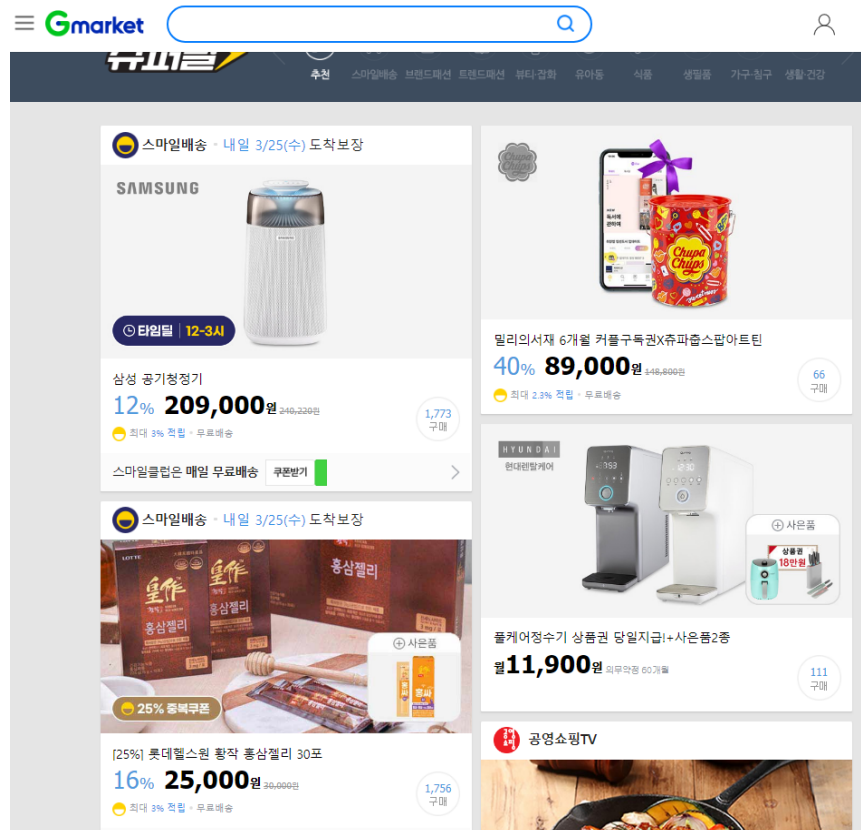
실제 강화학습이 어떻게 적용될 수 있을까?

2.1 주제

- 쇼핑몰에서 어떤 콘텐츠/어떤 상품을 상위에 보여줄수있을까?
- 특정 카테고리내에서 어떤 상품을 상위에 노출시킬까?
- 소비자에게 어떤 광고를 보여줄수있을까?
- 소비자에게 어떤 상품을 추천상품으로 보여줄 수 있을까?

결국 알고싶은 것은 어떠한 행동패턴을 보인 사용자의 정보가 어떠한지 와 각 특정 행동패턴의 사용자에게 어떤 상품을 추천할것인지/먼저 보여줄 것인지 입니다.

2.2 Environment : 쇼핑몰 상품 리스트 화면



- Environment: 쇼핑몰 상품 리스트 화면
 1. 쇼핑몰안에서도 다양한 리스트업 카테고리가 존재한다.
 1. 카테고리별 상품 페이지
 2. 할인 판매/ 추천상품 페이지
 3. 브랜드별 상품 페이지

2.3 State

- 상품 리스트 분야

- 상품리스트의 개수 : 한페이지에 상품이 몇개의 상품이 올라와 있는지, 일반적으로 상품은 한 페이지에서 최대 개수만큼 올라와 있으며 페이지로 나누기 때문에 **페이지의 개수도 적용가능**
- 광고상품의 개수: 상품페이지 안에서도 광고 상품을 올릴 수 있는데 한 페이지안에 광고 상품의 개수
- **시간 분야**
 - 상품 페이지의 특정위치(n Page)에서 머문시간(대기시간) : 소비자가 몇페이지에 머물고 있는지, 가장 오래 머문 페이지는 어디지(1페이지 제외)
 - 특정위치에서 사람들의 평균 대기시간 : 한페이지 안에서도 스크롤의 위치에 따라 상,중,하로 나누어 각 위치별로 평균 머무는 시간
 - 특정위치에서 가장 오래 머문시간 : 단순 대기상황이 아니라면, 상품에 대한 고민을 하고 있는 시간으로 파악해 볼 수 있음
- **회원 정보 분야**
 - 회원의 나이, 회원의 직업, 회원의 성별, 광고 클릭비율, 즐겨찾기여부, 장바구니 상품개수
 - 접속위치, 접속시간

2.4 Action : 상품 클릭

- Action : 특정 리스트에서 소비자가 상품페이지로 클릭하는 경우

2.5 Reward

Reward에 대한 개념이 부족하여 최대한 다양한 방식으로 Reward를 고민해 보았습니다.

- **Reward 1 : 전체 상품리스트의 길이/클릭한 상품의 순서**
 - 상위에 있는 상품리스트를 클릭할 수록 분모가 작아지기 때문에 보상이 커진다.

- **Rward 2 : 전체페이지의 길이/상품을 클릭한 페이지**
 - Reward 1의 경우에는 상품 리스트가 엄청나게 길 경우 점수가 너무 크기 때문에 페이지 단위로 Reward를 조절할 수 있음
- **Reward 3 : (평균적으로 리스트에서 상품을 클릭하는데 걸리는 시간) - (특정 소비자가 리스트에서 상품을 클릭하는데 걸리는 시간)**
 - 평균보다 빠르게 Action(클릭)을 할 경우 +점수를 받고 더욱 오래걸릴 경우 -점수를 받게 된다.
 - 상품을 클릭하는 것은 구매의도에 대한 확실한 정보를 얻기보다는 관심정도(빠르게 클릭할수록 호기심이 생기는/끌리는 상품일 것이다)를 보는 것이 때문

(Reward 설정이 가장 어려운 것 같습니다...)

3 적용 효과

(상상...)

- 경기도 수원에 거주하고 있는 20대의 대학생 여성이 쇼핑몰의 식품 카테고리에서 특정 체류시간이 길고, 클릭율이 높다. ⇒ **식품료에 대한 소비를 오프라인보다 온라인에서 할 가능성이 높은 회원으로 분류**
- 서울에 거주하고 있는 30대의 직장인 남성이 쇼핑몰의 전자기기 카테고리에서 체류시간이 길고 장바구니에는 평균적으로 10개 이상의 상품을 보관한다 ⇒ **전자기기에 관심이 많은 소비자, 장바구니의 평균 가격대가 높은 것으로 보아 고가의 상품을 복합적으로 고려하여 구매하는 회원으로 분류 → 전자기기 관련 할인 쿠폰 발송 고려**