

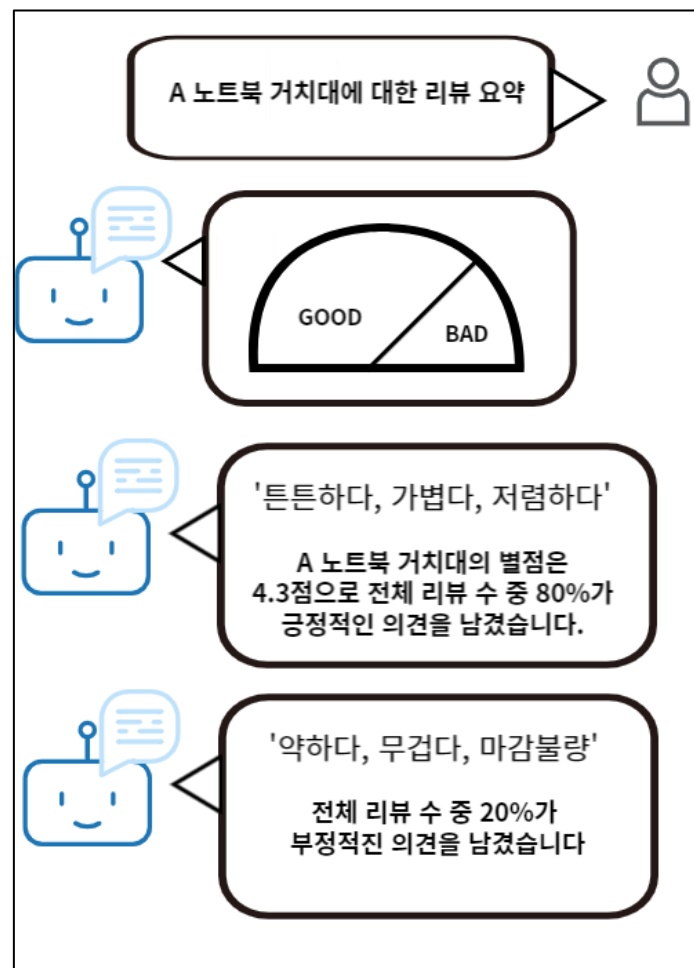
쇼핑몰 리뷰 분석 서비스(가제) 초기 개발 계획

구현해야 할 가장 큰 기능 2가지 => 요약, 추천

프로그램의 개발 방향성: 사용자에게 편리성 제공, 리뷰 검색에 소비되는 시간 단축

1. 요약 기능 구체화

- 한 상품에 대한 리뷰를 크롤링하고 리뷰 내용을 분석하여 긍정적, 부정적 리뷰를 분류한 뒤 가장 많이 나온 키워드를 도출한다. 도출된 키워드들을 이용해 **리뷰 요약**을 **문장형으로** 제공한다.
- 한 상품에 대한 긍정적, 부정적 리뷰 비율을 **수치로** 제공한다.



A. 왜 긍정, 부정 리뷰를 개수가 아닌 % 수치로 표현하는가?

- ✓ 한 리뷰 안에 긍정적, 부정적인 평가가 같이 있을 수 있기 때문에 리뷰를 개수로 제공하기 어렵다고 생각한다.

B. 긍정적, 부정적 리뷰에서 속성이 중복될 수도 있다. 이때 사용자 측면에서 요약본에 대한 신뢰도가 떨어질 수도 있지 않은가? 예) 물건이 튼튼해요, 물건이 약한 것 같아요.

- ✓ 소비자에 입장에서 긍정적 리뷰만이 아닌 부정적 리뷰에도 관심을 갖게 될 수 있기 때문에 같은 속성이라고 할 지라도 긍정적, 부정적인 리뷰를 분리하여 제공해야한다. 또한 한 상품에 대한 **전체적인 리뷰의 긍정적, 부정적 리뷰 비율을** 제공하여 사용자에게 자율적인 판단을 맡긴다.

예) 물건이 튼튼해요(80%), 물건이 약해요(20%)

2. 추천기능의 개발 계획 (구체화)

- 사용자의 검색 기록을 바탕으로 상품을 추천한다.
- 사용자가 검색한 상품의 같은 카테고리 내의 다른 상품을 추천한다.
- 상품의 추천과 함께 추천 이유를 제공한다.

문제1) 추천 제공 형식은 어떻게 할 것인가?

- ✓ 별도의 화면/광고와 같은 배너 형태 등 (미정).

문제2) 검색 결과를 바탕으로 같은 카테고리 내에 다른 상품을 추천할 때 사용자가 그 추천 항목에 대해 만족감을 느낄까?

예) 초기 노트북 검색결과가 만족스럽지 않았을 때 다른 노트북 상품을 추천 하는게 아닌 노트북 주변기기들을 추천하게 되는 경우.

- 의견 1) 추천을 제공할 때 부정적인 리뷰 분석하여 상품의 문제점을 보완할 수 있는 유사 상품이나 동일한 카테고리 내에 다른 상품을 추천해 주자.
- 의견 2) 상품을 검색하고 검색한 데이터가 쌓이면 쌓인 데이터를 이용해서 상품을 추천 해주자. (A와 B를 검색했을 때, A와 B의 리뷰를 분석하여 추천하자)
- 의견 3) 사용자가 검색한 상품과 유사한 상품을 추천하되 사용자가 원할 때 같은 카테고리 내에서 다른 상품들을 추천하자.
- 의견 4) 긍정적인 리뷰와 유사 상품을 제공하고 부정적인 리뷰와 보완할 만한 상품을 추천하자.

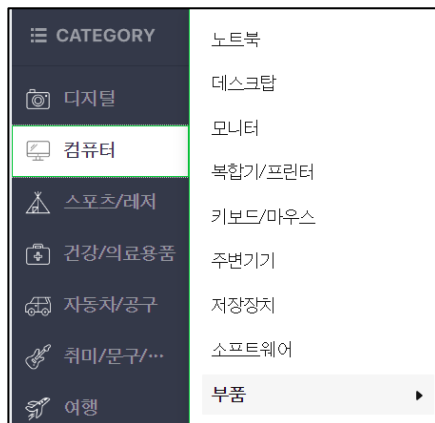
3. 카테고리 결정

여성패션 / 남성패션 / 화장품, 미용 / **가구, 인테리어** / 식품 / 생활, 주방용품 / 가전, 디지털 /

컴퓨터 / 스포츠, 레저/ 건강, 의료용품/ 취미, 문구, 악기

문제) 예시로 컴퓨터 카테고리 내에 여러가지 카테고리들이 있는데 나뉜진 카테고리들도 규모가 크기 때문에 컴퓨터 카테고리 안에서 다른 카테고리를 골라도 되는가?

예) 노트북, 데스크탑, 모니터, 키보드/마우스



4. 크롤링 테스트

테스트 진행 내용: 네이버에서 제공하는 검색 API는 네이버 쇼핑에서 검색한 물품 목록에 대한 결과만을 가져오는 한계가 있기 때문에 네이버 검색 API와 파이썬 REQUEST, URLLIB 라이브러리를 사용하여 테스트를 진행하였다. 테스트 결과 리뷰 조회 조건에 대해 약 2000개의 리뷰까지 가져올 수 있었다.

전체 리뷰 중 2000개의 리뷰 만으로는 분석이 어렵다고 판단하여 '평점' 별로 2000개씩 리뷰 데이터를 추출하는 방법을 선택한다.

문제 1) 서버에서 제한을 거는 크롤링 요청 개수는 얼마인가? (추가 테스트 필요)

문제 2) 크롤링 시 얼마의 시간이 걸리는가?

- 2464개 = 30.7초 => **1초에 약 82개의 데이터 수집**