08-10

**2차 주제 회의**

~~1. 상품 추천 챗봇~~

~~1-2. 상품에 대한 정보 제공 (ex) LG gram 노트북 정보 🡪 LG gram은 oo 하다)~~

~~1-3. 상품(제품)에 대한 추천 (ex) oo 사양의 oo 노트북 추천해줘 🡪 1. LG gram 2. Apple)~~

1. ~~영화, 드라마 줄거리 학습시켜서 우리가 어떠한 줄거리를 입력하면 추천해주는 챗봇(박지혜님이 말씀하신 웹툰 추천과 비슷), 감독과 주연배우도 입력데이터로 사용~~
2. ~~기능에 맞는 운동화 추천(한수길님)~~

**2. 광고주를 위한 인플루언서 분석**

🡪 sns 데이터 가져오는게 힘들다는 단점(인스타그램만 가능)

🡪 유튜브, 네이버 블로그, 인스타그램 활용

인플루언서가 주로 광고하는 제품에 따라 분류하고 가짜 인플루언서를 구별(카테고리별 점수? 지수?를 제공)해주는 시스템?

제품에 맞는 인풀루언서를 추천해주고 선정에 필요한 점수(지수)를 제공, 가짜 인플루언서라 판별될 경우 위험성 점수(지수) 제공.

유튜버 구독자만 많은 사람(조회수가 낮은 사람) 🡪 구독자 대비 조회수

소통은 활발하지만 광고 조회수만 낮은 유튜버 🡪 조회수가 구독자수, 영상의 댓글 수 수집

일플루언서 판별을 위한 기준 선정 이유

기업 마케팅 시장에서 인터넷 인플루언서(influencer) 영향력이 커지면서 이를 활용한 기업도 증가하고 있다.

- 인플루언서 마케팅의 최고 장점은 특정 카테고리에 관심이 있는 타겟층에게 마케팅을 진행할 수 있는 점에 있다.

'소셜미디어 블랙마켓'

- 인위적으로 팔로워 수를 늘려주는 매크로 프로그램이 팔리는 등 인터넷 인플루언서의 실제 영향력은 과장되었거나 심지어 조작하는 시장이 형성되고 있음.

1. 보통 정상적인 팔로워 증가는 자연스럽게 우상향을 보이고 있지만, 비정상적인 증가하는 경우 계단식으로 특정기간에 증가하는 경우가 있다.

- 특정 시간대에 소셜미디어 블랙마켓에서 거래된 매크로로 팔로워가 비정상적인 증가할 수 있음.

2. 정상적인 인플루언서(방송 연예인, 분야 저명인사 등은 제외)로 알려진 사람들의 계정은 보통 팔로워 대비 게시물당 평균 좋아요는 1~3% 내외이다.

- 팔로워 수가 1만명인데 좋아요 수가 게시물 하나당 1000개 이상(10%이상)이면 의심해볼 수 있다.

3. 소셜미디어 블랙마켓에서 외국인 팔로워는 비교적 손쉽고 저렴하게 늘릴 수 있다. 국내 인플루언서의 팔로워 상당수가 외국인이라면 의심해 볼 필요가 있다.

4. 팔로워를 구매한 경우 팔로워 숫자는 높아 보이나, 유령 계정인 가짜 팔로워 들은 인플루언서의 콘텐츠에 참여하지 않고, 비정상적으로 낮은 참여율을 갖음. 반대로 인위적으로 좋아요, 댓글 등의 서비스를 구매하여 참여 수준을 높이는 인플루언서는 매우 높은 참여율을 갖음.

- 참여율 계산은 [게시물 참여 수/팔로워 수 \* 100]으로 계산한다.

5. 낮은 품질의 댓글 참여

- 게시물과 관련 없는 댓글이 많을수록, 내용이 없는 이모티콘이나 댓글이 많은 경우 가짜 인플루언서일 가능성이 높다.

**광고주를 위한 인플루언서 분석**

가짜 인플루언서 지도 학습으로 구별하는 건 좋은 생각이지만 데이터를 크롤링 하는데 어렵다.

네이버 인플루언서에 한정 지어서 분석 및 분류. 광고 효과 분석.

네이버 인플루언서의 경우 유튜브와 인스타그램 계정까지 연동하여 정보 제공하고 있음.

네이버 인플루언서가 되기 위해서는 아무나 지원할 수 있지만 네이버의 심사를 거쳐야 가능 🡪 그러면 돈 주고 인플루언서가 될 수 없는 걸까? 🡪 암암리에 해주는 업체가 있다고는 함 🡪 네이버 측에서도 주기적으로 감시를 함

**지도 학습을 하기 위한 데이터 수집**

1. 팔로워 증가 추이를 분석

🡪 인스타그램이나 네이버에서 제공하지 않아서 자료를 찾기 어려움.

2. 댓글 분석을 통해 가짜 인플루언서 판별

3. 참여율 계산

🡪 현재 팔로워 수로 해야 할지 고민이 필요.

4. 해외 팔로워 비율을 파악해서 비율이 너무 비정상인지 판별

🡪 크롤링으로 파악이 가능 한가???

5. 네이버 인플루언서의 경우 본문 내용으로 구분을 시킬 수 있을까?

지도 학습 🡪 회귀분석?? 분류?? 이진분류?? 다중분류??  
이진 분류를 통해서 확률을 점수로 제공?

**사각화 자료만 제공 하게 된다면?**

시각화 자료를 제공 🡪 기존 인플루언서 알선? 추천?이랑 비슷

지도 학습을 사용하여 인플루언서의 이미지나 주로 홍보하는 상품(제품) 분류 🡪 광고주에 어울리는 인플루언서를 추천 🡪 점수나 시각화 자료를 통해 유령 또는 가짜 인플루언서인지를 판단할 수 있게 제공

🡪 보고서 형식!

🡪 일반 유료로 제공해주는 사이트와 비슷하게 되지 않을까 합니다.

**네이버 블로그 인플루언서에만 초첨을 맞춰서 정보를 제공한다면??**

네이버 블로그 인플루언서로 한다면?

🡪 이 사람이 진짜 영향력이 있는 인플루언서인지, 아니면 매일 가십거리 글만 써서 단순히 일방문자수만 높고 영향력은 없는 사람인건지?

🡪 본문 크롤링하여 지도학습 돌릴 수 있지 않을까요?

🡪 정답데이터는 이진분류로 한다면?

🡪 일 방문자수 / 구독? 이웃? / 좋아요(공감) 비율 / 댓글 수와 질을 파악 하여 영향력이 있는지 판단

🡪

인스타그램 문제점

1. 1만~10만 팔로워만 특정지어서 검색이 불가능(다른 제공하는 사이트가 있지만 유료사이트)

2. 인스타그램에서 제공하는 데이터가 별로 없음.(팔로워 추이 등등)

우리가 인스타그램에서 크롤링 해올 수 있는 것

1. 본문 내용

2. 댓글

3. 좋아요

4. 해시태그

5. 장소